REMARKS

Applicants reserve the right to prosecute non-elected subject matter in subsequent divisional applications.

Claims 3-13 and 30-45 have been canceled. Claims 1, 2, 14-29, 46, and 47 are currently pending in the present Application. Claims 1, 2, 17, 18 and 46 are actively being prosecuted, having been elected with traverse in the response to Restriction Requirement filed October 29, 2002. Please note that claims 1, 2, 18, 25 and 26 were amended in the response to Restriction Requirement filed October 29, 2002. Applicants request acknowledgment by the Examiner in writing that said amendments were entered into the record of the instant Application.

Rejoinder of Claims

Applicants continue to request the rejoinder of claims 19, 20, 23 and 26-28 which are "method of making" and "method of use" for the polypeptides of product claim 1. Therefore, upon allowance of a product claim, it is believed that claims 19, 20, 23 and 26-28 should be rejoined and considered in accordance with the Commissioner's Notice in the Official Gazette of March 26, 1996, entitled "Guidance on Treatment of Product and Process Claims in light of *In re Ochiai, In re Brouwer* and 35 U.S.C. § 103(b)." See also MPEP § 821.04 Rejoinder which states:

> if applicant elects claims directed to the product, and a product claim is subsequently found allowable, withdrawn process claims which depend from or otherwise include all the limitations of the allowable product claims will be rejoined.

Amendment to the Claims

Claim 1 has been amended to include in claims 1 b) and 1 c) recitation of, "said polypeptide having DNA binding activity." Support for this amendment can be found throughout the Specification. See for example, page 4, lines 7-10, (Applicants identify SEQ ID NO:1 as a prostate associated Ets protein), page 1, lines 15-17, page 2, lines 1-2 and 12-13 (Ets proteins have a characteristic Ets domain which binds DNA) and page 17, lines18-27 (Ets domains identified in SEQ ID NO:1). These amendments are being made in order to expedite prosecution and not for reasons related to patentability. Therefore, Applicants respectfully request entry of these amendments.

Objection of Claims 2, 18 and 46

Claims 2, 18 and 46 are objected to for being dependent upon rejected claims. Applicants request said objections be held in abeyance until allowable subject matter has been identified.

Rejection under 35 U.S.C. §112, first paragraph, written description

Claims 1 and 17 were rejected under 35 U.S.C. §112, first paragraph, allegedly because the Specification contained "subject matter which was not described in the specification in such a way as to reasonably convey to one skilled in the relevant art that the inventor(s), at the time the application was filed, had possession of the claimed invention." The Office Action asserts that:

- There is insufficient written description to show that Applicant was in possession of a polypeptide that is 90% identical to SEQ ID NO. 1. . . The full scope in Applicant's invention encompasses making any one of 20 changes to any one of 33 amino acids (20 known amino acids and 90% identity would allow one to change any one of up to 33 amino acids in a 335 amino acid protein). 33^{20} changes leads the full scope of Applicant claimed invention to be upto [sic] [33^{20}] polypeptides. Does Applicant contend that the written description of one sequence describes a genus of 23457341881036792870784632736O1 polypeptides, just because they have claimed the sequence? The Examiner disagrees. (Office Action of July 14, 2003, page 2);

- In addition the term "biologically active fragment or immunologically active fragment" would include an essentially unlimited number of undefined compounds. (Office Action of July 14, 2003, page 2);

- For inventions in an unpredictable art, adequate written description of a genus which embraces widely variant species *cannot* be achieved by disclosing only one species within the genus. (Office Action of July 14, 2003, page 3); [emphasis in original]

- Applicant has no written support for the term naturally in the specification or claims as originally filed. (Office Action of July 14, 2003, page 3)

These rejections are respectfully traversed.

The requirements necessary to fulfill the written description requirement of 35 U.S.C. 112, first paragraph, are well established by case law.

> . . . the applicant must also convey with reasonable clarity to those skilled in the art that, as of the filing date sought, he or she was in possession *of the invention.* The invention is, for purposes of the "written description" inquiry, *whatever is now claimed. Vas-Cath, Inc. v. Mahurkar*, 19 USPQ2d 1111, 1117 (Fed. Cir. 1991)

> . . . Mention of representative compounds encompassed by generic claim language ***clearly is not required by Section 112 or any other provision of the statute***. But, where no explicit description of a generic invention is to be found in the specification...mention of representative compounds may provide an implicit description upon which to base generic claim language. *In re Robins*, 429 F.2d 452, 456-57, 166 USPQ 552, 555 (CCPA 1970) [emphasis added]

> . . . [I]t has been consistently held that the naming of one member of such a group is not, in itself, a proper basis for a claim to the entire group. However, ***it may not be necessary to enumerate a plurality of species if a genus is sufficiently identified in an application by 'other appropriate language.'*** *In re Grimme*, 274 F.2d 949, 952, 124 USPQ 499, 501 (CCPA 1960) [emphasis added]

Attention is also drawn to the Patent and Trademark Office's own "Guidelines for Examination of Patent Applications Under the 35 U.S.C. Sec. 112, para. 1", published January 5, 2001, which provide that:

> An applicant may also show that an invention is complete by disclosure of sufficiently detailed, relevant identifying characteristics which provide evidence that applicant was in possession of the claimed invention, i.e., ***complete or partial structure***, other physical and/or chemical properties, functional characteristics when coupled with a known or disclosed correlation between function and structure, or some combination of such characteristics. What is conventional or well known to one of ordinary skill in the art need not be disclosed in detail. ***If a skilled artisan would have understood the inventor to be in possession of the claimed invention at the time of filing, even if every nuance of the claims is not explicitly described in the specification, then the adequate description requirement is met.*** [footnotes omitted, emphasis added]

Thus, the written description standard is fulfilled by both what is specifically disclosed and what is conventional or well known to one skilled in the art.

A. The Specification Provides an Adequate Written Description of the Claimed "Variants" of SEQ ID NO:1

SEQ ID NO:1 is specifically disclosed in the application (see, for example, page 4, lines 7-11, page 17, line 12 to page 18, line 10 and Figures 1A,1B, 1C, 1D, and 1E). Polypeptide variants having at least 90% identity to SEQ ID NO:1 are described, for example, at page 18, lines 15-18. Moreover, the Specification at page 7, lines 17-20, defines "PRAEP" as "the amino acid sequences of substantially purified PRAEP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and preferably the human species, from any source, whether natural, synthetic, semi-synthetic, or recombinant." Hence, by referring to "a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence of SEQ ID NO:1, said polypeptide having DNA binding activity" it is apparent that the inventors contemplated naturally occurring variants of SEQ ID NO:1 as opposed to variants created *in vitro*, i.e., those variants known to one of skill in the art to be synthesized, semi-synthesized or produced by recombinant methodologies.

Additionally, the term "naturally occurring" is a well-known term in the art which Applicants intended to be used in such context. As such, no further definition of the term is necessary (MPEP 2163 IIA3(a)):

> What is conventional or well known to one of ordinary skill in the art need not be disclosed in detail. See *Hybritech Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d at 1384, 231 USPQ at 94. If a skilled artisan would have understood the inventor to be in possession of the claimed invention at the time of filing, even if every nuance of the claims is not explicitly described in the specification, then the adequate description requirement is met. See, e.g., *Vas-Cath*, 935 F.2d at 1563, 19 USPQ2d at 1116; *Martin v. Johnson*, 454 F.2d 746, 751, 172 USPQ 391, 395 (CCPA 1972) (stating "the description need not be in *ipsis verbis* [i.e., "in the same words"] to be sufficient").

One of ordinary skill in the art would recognize that "***a naturally occurring amino acid sequence***" as recited in claim 1 is one which occurs in nature. Through the process of natural selection, nature will have determined the appropriate amino acid sequences. Given the information provided by SEQ ID NO:1 (the amino acid sequence of PRAEP) and SEQ ID NO:2 (the polynucleotide sequence encoding PRAEP), one of skill in the art would be able to routinely obtain "a naturally-occurring amino acid sequence at least 90% identical to the amino acid sequence of SEQ ID NO:1, said polypeptide

having DNA binding activity." For example, the identification of relevant polynucleotides could be performed by hybridization and/or PCR techniques that were well-known to those skilled in the art at the time the subject application was filed and/or described throughout the Specification of the instant application. For example:

> The invention further provides an isolated and purified polynucleotide which hybridizes under stringent conditions to the polynucleotide encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1, as well as an isolated and purified polynucleotide which is complementary to the polynucleotide encoding the polypeptide comprising the amino acid sequence of SEQ ID NO:1 or a fragment of SEQ ID NO:1. (Specification at page 4, lines 19-24.)
>
> The term "stringent conditions," refers to conditions which permit hybridization between polynucleotides and the claimed polynucleotides. Stringent conditions can be defined by salt concentration, the concentration of organic solvent (e.g., formamide), temperature, and other conditions well known in the art. In particular, stringency can be increased by reducing the concentration of salt, increasing the concentration of formamide, or raising the hybridization temperature. (Specification at page 14, line 29 to page 15, line 4.)
>
> Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:2, or a fragment of SEQ ID NO:2, under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol. 152:507-511.) (Specification at page 19, lines 26 to 30.)
>
> In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding PRAEP or closely related molecules may be used to identify nucleic acid sequences which encode PRAEP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only naturally occurring sequences encoding PRAEP, allelic variants, or related sequences.
>
> Probes may also be used for the detection of related sequences, and should preferably have at least 50% sequence identity to any of the PRAEP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:2 or from genomic sequences including promoters, enhancers, and introns of the PRAEP gene. (Specification at page 40, lines 13 to 25.)

See also Example VI at pages 52-53.

Thus, the Examiner's assertion that "[t]he full scope in Applicant's invention encompasses making any one of 20 changes to any one of 33 amino acids...up to 2345734188103679287078463273601 polypeptides" is clearly in error (See Office Action mailed July 14, 2003 at page 2). In order to make the claimed polypeptides, one skilled in the art need only screen a cDNA library or use appropriate PCR conditions to identify relevant polynucleotides/polypeptides that already exist in nature. By adjusting the nature of the probe or nucleic acid (*i.e.*, non-conserved, conserved or highly conserved) and the conditions of hybridization (maximum, high, intermediate or low stringency), one can obtain variant polynucleotides of SEQ ID NO:2 which, in turn, will allow one to make the variant polypeptides of SEQ ID NO:1 recited by the present claims. Conventional methods for making polypeptides, such as those methods described at page 22, lines 5-11 of the Specification, could be used to make the recited polypeptide variants.

Thus, not only can one of ordinary skill in the art based on hybridization results, predict, detect and isolate with a high level of certainty, the presence or absence of a polynucleotide sequence encoding a naturally occurring amino acid sequence at least 90% identical to the amino acid sequence of SEQ ID NO:1, said polypeptide having DNA binding activity, the skilled artisan is also able to compare the encoded polypeptide sequence to SEQ ID NO:1 to determine the percentage identity of said polypeptide sequence to SEQ ID NO:1. Such methods for isolating, cloning, expressing and comparing polynucleotides and their encoded polypeptides are routine analyses for the skilled artisan in the field of recombinant molecular biology.

Accordingly, the Specification provides an adequate written description of the recited variants.

1. The present claims specifically define the claimed genus through the recitation of chemical structure

Court cases in which "DNA claims" have been at issue (which are hence relevant to claims to proteins encoded by the DNA) commonly emphasize that the recitation of structural features or chemical or physical properties are important factors to consider in a written description analysis of such claims. For example, in *Fiers v. Revel*, 25 USPQ2d 1601, 1606 (Fed. Cir. 1993), the court stated

that:

> If a conception of a DNA requires a precise definition, such as by structure, formula, chemical name or physical properties, as we have held, then a description also requires that degree of specificity.

In a number of instances in which claims to DNA have been found invalid, the courts have noted that the claims attempted to define the claimed DNA in terms of functional characteristics without any reference to structural features. As set forth by the court in *University of California v. Eli Lilly and Co.*, 43 USPQ2d 1398, 1406 (Fed. Cir. 1997):

> In claims to genetic material, however, a generic statement such as "vertebrate insulin cDNA" or "mammalian insulin cDNA," without more, is not an adequate written description of the genus because it does not distinguish the claimed genus from others, except by function.

Thus, the mere recitation of functional characteristics of a DNA, without the definition of structural features, has been a common basis by which courts have found invalid claims to DNA. For example, in *Lilly*, 43 USPQ2d at 1407, the court found invalid for violation of the written description requirement the following claim of U.S. Patent No. 4,652,525:

> 1. A recombinant plasmid replicable in procaryotic host containing within its nucleotide sequence a subsequence having the structure of the reverse transcript of an mRNA of a vertebrate, which mRNA encodes insulin.

In *Fiers*, 25 USPQ2d at 1603, the parties were in an interference involving the following count:

> A DNA which consists essentially of a DNA which codes for a human fibroblast interferon-beta polypeptide.

Party Revel in the *Fiers* case argued that its foreign priority application contained an adequate written description of the DNA of the count because that application mentioned a potential method for isolating the DNA. The Revel priority application, however, did not have a description of any particular DNA structure corresponding to the DNA of the count. The court therefore found that the Revel priority application lacked an adequate written description of the subject matter of the count. Thus, in *Lilly* and *Fiers*, nucleic acids were defined on the basis of functional characteristics and were found not to comply with the written description requirement of 35 U.S.C. §112; *i.e.*, "an mRNA of a vertebrate, which mRNA encodes insulin" in *Lilly*, and "DNA which codes for a human fibroblast interferon-beta polypeptide" in *Fiers*. In contrast to the situation in *Lilly* and *Fiers*, the claims at issue

in the present application define polypeptides of SEQ ID NO:1 in terms of chemical structure, rather than in terms of functional characteristics. For example, the "variant language" of independent claim 1 recites chemical structure to define the claimed genus:

> 1. An isolated polypeptide sequence selected from the group consisting of ...b) a polypeptide comprising a naturally occurring amino acid sequence at least 90% identical to an amino acid sequence of SEQ ID NO:1, said polypeptide having DNA binding activity, ...

From the above it should be apparent that the claims of the subject application are fundamentally different from those found invalid in *Lilly* and *Fiers*. The subject matter of the present claims is defined in terms of the chemical structure of SEQ ID NO:1. In the present case, there is no reliance merely on a description of functional characteristics of the claimed polypeptides. Moreover, the functional recitation, as amended, adds to the structural characterization of the claimed polypeptides. The polypeptides defined in the claims of the present application recite structural features and functional characteristics of the polypeptide of SEQ ID NO:1, and cases such as *Lilly* and *Fiers* stress that the recitation of structure is an important factor to consider in a written description analysis of claims of this type. By failing to base its written description inquiry "on whatever is now claimed," the Examiner failed to provide an appropriate analysis of the present claims and how they differ from those found not to satisfy the written description requirement in *Lilly* and *Fiers*

2. The present claims do not define a genus which is "highly variant"

Furthermore, the claims at issue do not describe a genus which could be characterized as "highly variant." Available evidence illustrates that the claimed genus is of narrow scope. In support of this assertion, the Examiner's attention is directed to the reference by Brenner et al. ("Assessing sequence comparison methods with reliable structurally identified distant evolutionary relationships," Proc. Natl. Acad. Sci. USA (1998) 95:6073-6078), enclosed herewith. Through exhaustive analysis of a data set of proteins with known structural and functional relationships and with <40% overall sequence identity, Brenner et al. have determined that 30% identity is a reliable threshold for establishing evolutionary homology between two sequences aligned over at least 150 residues. (Brenner et al., pages 6073 and 6076.) Furthermore, local identity is particularly important in this case for assessing the significance of the alignments, as Brenner et al. further report that ≥40% identity over

at least 70 residues is reliable in signifying homology between proteins. (Brenner et al., page 6076.) The present application is directed, *inter alia*, to Ets proteins related to the amino acid sequence of SEQ ID NO:1. In accordance with Brenner et al, naturally occurring molecules may exist which could be characterized as Ets proteins and which have as little as 30% identity over at least 150 residues to SEQ ID NO:1. The "variant language" of the present claims recites, for example, an isolated polypeptide comprising an amino acid sequence that is "a naturally occurring amino acid sequence having at least 90% sequence identity to the sequence of SEQ ID NO:1, said polypeptide having DNA binding activity" (note that SEQ ID NO:1 has 335 amino acid residues). This variation is far less than that of all potential Ets proteins related to SEQ ID NO:1, i.e., those Ets proteins having as little as 30% identity over at least 150 residues to SEQ ID NO:1. Moreover, such variation, either a polypeptide having an amino acid sequence having at least 90% identity to an amino acid sequence of SEQ ID NO:1, said polypeptide having DNA binding activity or a biologically active fragment of a polypeptide having an amino acid sequence of SEQ ID NO:1, said fragment having DNA binding activity, is orders of magnitude less that the 33^{20} potential polypeptides alleged by the Examiner. An assertion otherwise ignores that such variants and biologically active fragments are ***naturally*** occurring.

3. Ets proteins contain Ets domains which are known to function in DNA binding

Applicants have adequately described the instant invention in terms of a method of its making together with providing both structural and functional characteristics and have provided evidence of a correlation between structure and function. Conventional methods for making the claimed polypeptides, are described at page 22, lines 5-11 of the Specification. Applicants have identified SEQ ID NO:1 as a prostate associated Ets protein (Specification, page 4, lines 7-10). The Specification teaches that Ets proteins have a characteristic Ets domain which binds DNA (Specification, page 1, lines 15-17; page 2, lines 1-2 and lines 12-13) and the Specification identifies the presence of Ets domains within SEQ ID NO:1 (Specification, page 17, lines18-27). Thus, SEQ ID NO:1, having ETS domains, would be understood to function in DNA binding.

The MPEP states that:

> The claimed invention as a whole may not be adequately described where an invention is described solely in terms of a method of its making coupled with its function and there is no described or art-recognized correlation or relationship between the structure of the invention and its function. (MPEP § 2163 I. A, Rev. 1, Feb. 2003, p. 2100-161).

In contrast, the instant invention is adequately described as claimed and is described in terms of the correlation of structure and function, as one of skill in the art would recognize that Ets domains correlate with DNA binding.

In support of this position, Applicants respectfully bring to the attention of the Examiner the publication of Boulukos, K.E. et al., (K.E. Boulukos et al., (1989) Mol. Cell Biol. 9:5718-21, Attachment B; abstract enclosed herewith). Boulukos et al. report that deletion analysis of the Ets domains indicates that the Ets domains are essential for integrity of the carboxy-terminal domain, and are essential for nuclear targeting and DNA binding activity *in vitro*. These results are further substantiated for SEQ ID NO:1 by Oettgen,P. et al., (Oettgen, P. et al., (2000) J. Biol. Chem. 275:1216-1225, Attachment C; enclosed herewith). Oettgen teaches that SEQ ID NO:1 (SEQ ID NO:1 is 100% identical to PDEF, see sequence alignment of SEQ ID NO:1 to g4007418, Attachment D) binds with high affinity to the GGAT-containing oligonubleotide of the PSA promoter E site, and conclude that SEQ ID NO:1 (PDEF) DNA binding specificity is distinct from other members of the Ets family with a unique preference toward the GGAT recognition sequence (Oettgen, et al., pp. 1221-2). Thus, Applicants have provided an art-recognized correlation between the Ets domain, a structural domain within the claimed invention, and the function of the Ets domain in DNA binding activity.

One of skill in the art would conclude that in order to have DNA binding activity, 90% variants and biologically active fragments of SEQ ID NO:1 would have the Ets domain as a structural domain of 90% variants and biologically active fragments of SEQ ID NO:1 in order to maintain DNA binding activity. Therefore, one of ordinary skill in the art would conclude that Applicants have provided an adequate description of SEQ ID NO:1 and variants thereof as well as biologically active fragments of SEQ ID NO:1 in terms of the correlation of the Ets domain, a structural component of SEQ ID NO:1, 90% variants and biologically active fragments thereof which are involved in the DNA binding activity of the claimed invention.

4. The state of the art at the time of the present invention is further advanced than at the time of the *Lilly* and *Fiers* applications

In the *Lilly* case, claims of U.S. Patent No. 4,652,525 were found invalid for failing to comply with the written description requirement of 35 U.S.C. §112. The '525 patent claimed the benefit of priority of two applications, Application Serial No. 801,343 filed May 27, 1977, and Application Serial No. 805,023 filed June 9, 1977. In the *Fiers* case, party Revel claimed the benefit of priority of an

Israeli application filed on November 21, 1979. Thus, the written description inquiry in those case was based on the state of the art at essentially at the "dark ages" of recombinant DNA technology.

The present application has a priority date of 04 April 1998. Much has happened in the development of recombinant DNA technology in the 17 or more years from the time of filing of the applications involved in *Lilly* and *Fiers* and the present application. For example, the technique of polymerase chain reaction (PCR) was invented. Highly efficient cloning and DNA sequencing technology has been developed. Large databases of protein and nucleotide sequences have been compiled. Much of the raw material of the human and other genomes has been sequenced. With these remarkable advances one of skill in the art would recognize that, given the sequence information of SEQ ID NO:1 and the additional extensive detail provided by the subject application, the present inventors were in possession of the polypeptide variants recited by the claims at the time of filing of this application.

5. The term "naturally" is fully supported in the claims and the Specification of the Application as originally filed

In the present case, the application as originally filed discloses "90% sequence variants" of SEQ ID NO:1 (PRAEP) (see e.g., page 18, lines 15-18) as well as naturally occurring polynucleotide sequences that encode such variants. Section (b) of claim 1 as originally filed recites ". . . a ***naturally*** occurring amino acid sequence at least 90% identical to an amino acid sequence of SEQ ID NO:1 . . ." See page 63 of the application as originally filed. As decided in *In re* Gardner, the claims itself may describe the disclosed invention even if the application has no corresponding disclosure (*In re* Gardner, 178 USPQ 149 (C.C.P.A. 1973). Hence, contrary to the Examiner's assertion otherwise, the term "naturally" is fully supported in claim 1 as it originally appeared.

Moreover, the Specification is replete with references to naturally occurring polynucleotides/polypeptides. For example:

> Where "amino acid sequence" is recited herein to refer to an amino acid sequence of a ***naturally*** occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule. (Specification at page 8, lines 24-27, emphasis added.)

The words "insertion" or "addition," as used herein, refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively, to the sequence found in the ***naturally*** occurring molecule. (Specification at page 13, lines 5-7, emphasis added.)

As used herein, the term "biologically active," refers to a protein having structural, regulatory, or biochemical functions of a ***naturally*** occurring molecule. . . . (Specification at page 9 lines 29-30, emphasis added.)

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding PRAEP, some bearing minimal similarity to the polynucleotide sequences of any known and ***naturally*** occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of ***naturally*** occurring PRAEP, and all such variations are to be considered as being specifically disclosed. (Specification at page 19, lines 1-8, emphasis added.)

Although nucleotide sequences which encode PRAEP and its variants are preferably capable of hybridizing to the nucleotide sequence of the ***naturally*** occurring PRAEP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding PRAEP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding PRAEP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the ***naturally*** occurring sequence. (Specification at page 19, lines 9-19, emphasis added.)

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to PRAEP have an amino acid sequence consisting of at least about 5 amino acids, and, more preferably, of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein and contain the entire amino acid sequence of a small, ***naturally*** occurring molecule. Short stretches of PRAEP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced. (Specification at page 32, lines 3-9, emphasis added.)

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding PRAEP or closely related molecules may be used to identify nucleic acid sequences which encode PRAEP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification (maximal, high, intermediate, or low), will determine whether the probe identifies only ***naturally*** occurring sequences encoding PRAEP, allelic variants, or related sequences. (Specification at page 40, lines 13-20, emphasis added.)

In another embodiment of the invention, nucleic acid sequences encoding PRAEP may be used to generate hybridization probes useful in mapping the ***naturally*** occurring genomic sequence. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) (Specification at page 44, lines 19-25, emphasis added.)

Sequences complementary to the PRAEP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of ***naturally*** occurring PRAEP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 primer analysis software and the coding sequence of PRAEP. To inhibit transcription, a complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the PRAEP-encoding transcript. (Specification at Example VIII at page 53, line 27 to page 54, line 6, emphasis added.)

Naturally occurring or recombinant PRAEP is substantially purified by immunoaffinity chromatography using antibodies specific for PRAEP. An immunoaffinity column is constructed by covalently coupling anti-PRAEP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Pharmacia). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing PRAEP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of PRAEP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/PRAEP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and PRAEP is

collected. (Specification at Example XIII at page 57, lines 15-25, emphasis added.)

In view of the foregoing evidence, Applicants submit that the rejection of claims 1 and 17 on the grounds that there is "no written support for the term [naturally] in the specification or claims as originally filed" is without merit, and is therefore improper.

Applicants are therefore confused by the Examiner's reference to MPEP section 2163.05(b), since the claims and the specification as originally filed EXPLICITLY supports the subgenus of naturally occurring 90% sequence variants of SEQ ID NO:1 as is now claimed in subsection b of claim 1. See e.g. the specification at page 18, lines 15-18 ". . . 90%, . . . sequence identity to the PRAEP amino acid sequence, and which contains at least one functional . . . characteristic of PRAEP."

6. Summary

The Office Action failed to base its written description inquiry "on whatever is now claimed." Consequently, the Action did not provide an appropriate analysis of the present claims and how they differ from those found not to satisfy the written description requirement in cases such as *Lilly* and *Fiers*. In particular, the claims of the subject application are fundamentally different from those found invalid in *Lilly* and *Fiers*. The subject matter of the present claims is defined in terms of the chemical structure of SEQ ID NO:1, including Ets domains known to one of skill in the art to function in DNA binding. The courts have stressed that structural features are important factors to consider in a written description analysis of claims to nucleic acids and proteins. In addition, the genus of polypeptides defined by the present claims is adequately described, as evidenced by Brenner et al. Additionally, 90% variants having DNA binding activity and biologically active fragments of SEQ ID NO:1 having DNA binding activity are adequately described, as evidenced by the Specification, and the teachings of Boulukos,K.E. et al. and Oettgen,P. et al. Furthermore, there have been remarkable advances in the state of the art since the *Lilly* and *Fiers* cases, and these advances were given no consideration whatsoever in the position set forth by the Office Action.

The Office Action failed to consider that the term "naturally" is a term that is well-known to those of skill in the art and, as such, does not require further definition in the Specification. For at least the above reasons it is believed that claims 1 and 17 meet the written description requirement of 35 U.S.C. § 112, first paragraph. It is therefore requested that this rejection be withdrawn.

Rejection under 35 U.S.C. §112, first paragraph, enablement

Claims 1 and 17 were rejected under 35 U.S.C. §112, first paragraph, allegedly because "the specification, while being enabling for the use of SEQ ID NO. 1, does not reasonably provide enablement for the use of any protein that is 90% identical to SEQ ID NO. 1 or any biological fragment of SEQ ID NO. 1 or immunological fragment of SEQ ID NO. 1" commensurate with the scope of the claimed invention. The Office further asserts that:

- . . . the present specification fails to disclose any protein which has ETS activity (Office Action of July 14, 2003, page 4).

- The total number of polypeptides encompassed by the claims is 2345734188103679287078463273601 polypeptides (Office Action of July 14, 2003, page 4).

Applicants travers this rejection for the reasons submitted below.

A. Legal Requirement

As set forth in *In re Marzocchi*, 169 USPQ 367, 369 (CCPA 1971):

> The first paragraph of § 112 requires nothing more than objective enablement. How such a teaching is set forth, either by the use of illustrative examples or by broad terminology, is of no importance.
>
> As a matter of Patent Office practice, then, a specification disclosure which contains a teaching of the manner and process of making and using the invention in terms which correspond in scope to those used in describing and defining the subject matter sought to be patented *must* be taken as in compliance with the enabling requirement of the first paragraph of § 112 *unless* there is reason to doubt the objective truth of the statements contained therein which must be relied on for enabling support.

To fulfill the enablement requirement of 35 U.S.C. §112, first paragraph, the claimed invention must be described in the Specification in such as way as to enable one skilled in the relevant art to which it pertains, or with which it is most nearly connected, to make and/or use the invention.

When determining whether the Specification meets the enablement requirement the courts have ruled that the claimed invention be disclosed in the patent together with information known in the art

such that one of ordinary skill in the art is *enabled* to make and use the invention without undue experimentation. See *In re Wands*, 858 F.2d at 737, 8 USPQ2d at 1404 (Fed. Cir. 1988); *United States v. Telectronics, Inc.*, 857 F.2d 778, 785, 8 USPQ2d 1217, 1223 (Fed. Cir. 1988). Thus, "a patent need not teach, and preferably omits, what is ***well known*** in the art. *In re Buchner*, 929 F.2d 660, 661, 18 USPQ2d 1331, 1332 (Fed. Cir. 1991); *Hybritech, Inc. v. Monoclonal Antibodies, Inc.*, 802 F.2d 1367, 1384, 231 USPQ 81, 94 (Fed. Cir. 1986), *cert. denied*, 480 U&.S. 947 (1987); and *Lindemann Maschinenfabrik GMBH v. American Hoist & Derrick Co.*, 730 F.2d 1452, 1463, 221 USPQ 481, 489 (Fed. Cir. 1984).

> In addition, the Manual of Patent Examination Procedure at § 2164.01(c) states: [A]s long as the specification discloses at least one method for making and using the claimed invention that bears a reasonable correlation to the entire scope of the claim, then the enablement requirement of 35 U.S.C. 112 is satisfied. *In re Fisher*, 427 F.2d 833, 839, 166 USPQ 18, 24 (CCPA 1970). Failure to disclose other methods by which the claimed invention may be made does not render a claim invalid under 35 U.S.C. 112. *Spectra-Physics, Inc. v. Coherent, Inc.*, 824 F.2d 1524, 1533, 3 USPQ2d 1737, 1743 (Fed. Cir.), *cert. denied*, 484 U.S. 954 (1987).

It is submitted that the Specification does reasonably provide an adequate written description to **enable** the claimed protein of SEQ ID NO:1 as well as the recited naturally-occurring 90% variants of "the amino acid sequence of SEQ ID NO:1, said polypeptide having DNA binding activity," biologically-active fragments "of a polypeptide having the amino acid sequence of SEQ ID NO:1, said polypeptide fragment having DNA binding activity" and "an immunogenic fragment of a polypeptide having an amino acid sequence of SEQ ID NO:1 as "now" claimed at the time of filing of this application.

The Examiner is well aware that the relative skill of those in the art is very high and the amount of direction or guidance needed to be disclosed in the Specification **to make** the protein of SEQ ID NO:1, and the recited "variants," "biologically-active fragments" and "immunogenic fragments" of SEQ ID NO:1 as "now" claimed is well within the grasp of one of skill in the art upon reading the specification, especially in light of Example VI (hybridization probes used to screen cDNAs, genomic DNAs or mRNAs for 90% variants of SEQ ID NO:1), discussed *supra*; Example X (assay for PRAEP activity, including biologically active fragments of SEQ ID NO:1); Example XII (production of anitbodies to PRAEP, including how to determine immunogenic fragments of SEQ ID NO:1); and those methods described at page 22, lines 5-11 for making the claimed polypeptides. Claimed variants of SEQ ID NO:1 are defined in the Specification at, for example, page 16, lines 22-30; and page 18, lines

15-18. Biologically active SEQ ID NO:1 is defined in the Specification at, for example, page 9 lines 29-30. Immunologically active SEQ ID NO:1 is defined in the Specification at, for example, page 9, line 30 to page 10, line 3.

Moreover, the functionality of the claimed variants will have been established anyway, since ***naturally occurring*** variants are claimed. Through the process of natural selection, nature will have determined the appropriate amino acid sequences. Given the information provided by SEQ ID NO:1 (the amino acid sequence of PRAEP) and SEQ ID NO:2 (the polynucleotide sequence encoding PRAEP), one of skill in the art would be able to routinely obtain: i) a polypeptide comprising an amino acid sequence of SEQ ID NO:1, ii) a polypeptide comprising a naturally-occurring amino acid sequence having at least 90% sequence identity to the sequence of SEQ ID NO:1, said polypeptide having DNA binding activity, iii) a biologically-active fragment of a polypeptide having the amino acid sequence of SEQ ID NO:1, said polypeptide having DNA binding activity, or iv) an immunogenic fragment of a polypeptide having an amino acid sequence of SEQ ID NO:1. As an additional example, the identification of relevant polynucleotides could be performed by hybridization and/or PCR techniques that were well-known to those skilled in the art at the time the subject application was filed and/or described throughout the Specification of the instant application. See, *e.g.*, page 40, lines 13 to 25 and Example VI at pages 52-53. Therefore, one skilled in the art need not make and test vast numbers of polypeptides that are based on the amino acid sequence of SEQ ID NO:1. Instead, one skilled in the art need only screen a cDNA library or use appropriate PCR conditions to identify relevant polynucleotides/polypeptides that already exist in nature.

Additionally, an assay for monitoring PRAEP activity is described in the Specification, for example, at page 55, Example X. Assays to determine functional activity are considered routine experimentation when identifying functional sequence variants. One of ordinary skill in the art would recognize polypeptide sequences which are variants having at least 90% amino acid identity to SEQ ID NO:1 or biologically active fragments of SEQ ID NO:1, as those polypeptides or fragments which, when assayed, have at least one function of a polypeptide comprising an amino acid sequence of SEQ ID NO:1. Accordingly, polypeptides comprising an amino acid sequence that is 90% identical to the amino acid sequence of SEQ ID NO:1 or biologically active fragments of SEQ ID NO:1 can easily be identified by one of skill in the art based on both the presence of functional and structural domains and by the assay, all disclosed in the Specification. Thus, one of skill in the art would understand upon reading the specification, how to make the protein of SEQ ID NO:1, and the recited "variants" and

"biologically-active fragments" of SEQ ID NO:1.

The Examiner attempts to provide further support for this rejection by citing Ngo et al. (in The Protein Folding Problem and Tertiary Structure Prediction, 1994, Birkhauser Boston, pages 433, 492-495). The Examiner asserts that "Ngo *et al* teach that "the relationship between the sequence of a peptide and its tertiary structure (i.e. its activity) are not well understood and are not predictable." (Office Action of July 14, 2003, at page 4).

It is not necessary to accurately predict protein function or structure and/or sequence in order to make and/or use the recited polypeptide variants, and biologically and immunologically active fragments. The Ngo et al. reference, which discusses methods of predicting protein structures is not relevant. The Ngo et al. reference has no bearing on the ability of a skilled artisan to screen a cDNA library or use appropriate PCR conditions to identify relevant polynucleotides, and their encoded polypeptides, that already exist in nature, without undue experimentation. Once again, the question is not whether the recited polypeptides retain the structure and/or function of the SEQ ID NO:1 polypeptide. The relevant question, for the purposes of enablement under 35 U.S.C. § 112, first paragraph, is whether a skilled artisan could make and use the claimed polypeptides. Regardless of whether a variant or fragment of SEQ ID NO:1 maintains the structure and/or function of the SEQ ID NO:1 polypeptide, that variant or fragment could still be used to make the polypeptides, without undue experimentation. Thus, the enablement requirement is satisfied.

In sum, contrary to the standard set forth in *Marzocchi, supra*, the Examiner has failed to provide any ***reasons*** why one would doubt that the guidance provided by the present Specification would enable one to make and use the protein of SEQ ID NO:1, and the recited "variants" and "biologically-active fragments" of SEQ ID NO:1. Hence, a *prima facie* case for non-enablement has not been established with respect to the protein of SEQ ID NO:1, and the recited "variants" and "biologically-active fragments" of SEQ ID NO:1.

Accordingly, for all the above reasons, the claimed subject matter is described in the Specification in such a way that one skilled in the art can make and/or use the claimed invention. Therefore, reconsideration and withdrawal of the rejection of claims 1 and 17 are respectfully requested.

CONCLUSION

In light of the above amendments and remarks, Applicants submit that the present application is fully in condition for allowance, and request that the Examiner withdraw the outstanding objections/rejections. Early notice to that effect is earnestly solicited.

If the Examiner contemplates other action, or if a telephone conference would expedite allowance of the claims, Applicants invite the Examiner to contact the undersigned at the number listed below.

Applicants believe that no fee is due with this communication. However, if the USPTO determines that a fee is due, the Commissioner is hereby authorized to charge Deposit Account No. **09-0108.**

Respectfully submitted,

INCYTE CORPORATION

Date: 14, October 2003

Shirley A. Recipon

Shirley A. Recipon
Reg. No. 47,016
Direct Dial Telephone: (650) 621-8555

Customer No.: 27904
3160 Porter Drive
Palo Alto, California 94304
Phone: (650) 855-0555
Fax: (650) 849-8886

Attachment(s):
A - Brenner et al., Proc. Natl. Acad. Sci. USA (1998) 95:6073-6078
B - K.E.Boulukos et al., (1989) Mol. Cell Biol. 9:5718-21
C - Oettgen, P. et al., (2000) J. Biol. Chem. 275:1216-1225
D - BLASTP of SEQ ID NO:1 verses Genpept137

Connective Tissue Growth Factor: a Cysteine-rich Mitogen Secreted by Human Vascular Endothelial Cells Is Related to the SRC-induced Immediate Early Gene Product CEF-10

Docket No.: PF-0501-1 DIV
USSN: 09/866,356
Attachment A

Douglass M. Bradham,* Atsuyuki Igarashi,* Robert L. Potter,§ and Gary R. Grotendorst*‡

Departments of *Ophthalmology, ‡Biochemistry and Molecular Biology, and §Chemistry, University of South Florida College of Medicine, Tampa, Florida 33620

Abstract. Human umbilical vein endothelial (HUVE) cells have been previously reported to express the genes for the A and B chains of PDGF and to secrete PDGF-related factors into culture media. Antihuman PDGF IgG affinity chromatography was used to purify PDGF-related activity from HUVE cell-conditioned media. Immunoblot analysis of the affinity-purified proteins with anti-PDGF IgG and antibodies specific for the A or B chain peptides of PDGF combined with chemotactic and mitogenic assays revealed that the major PDGF immunorelated molecule secreted by HUVE cells is a monomer of ~36–38 kD and that <10% of the purified biologically active molecules are PDGF A or B chain peptides. Screening of an HUVE cell cDNA library in the expression vector lambda gt11 with the anti-PDGF antibody resulted in the cloning and sequencing of a cDNA with an open reading frame encoding a 38-kD cysteine-rich secreted protein which we show to be the major PDGF-related mitogen secreted by human vascular endothelial cells. The protein has a 45% overall homology to the translation product of the v-src-induced CEF-10 mRNA from chick embryo fibroblasts. We have termed this new mitogen connective tissue growth factor.

THE vascular endothelium, which forms the nonthrombogenic lining of blood vessels, was once considered a relatively inert membrane. Recent studies of endothelial cells have demonstrated that they participate in numerous structural and physiological functions of the circulatory system. Structural proteins including collagens and laminin are secreted into the basement membrane, while procoagulant, anticoagulant, and platelet regulatory proteins and growth regulatory molecules are secreted from the luminal membrane into the circulating blood or into a wound site (Jaffe, 1984, 1987). Cultured endothelial cells have previously been shown to secrete factors that are chemotactic and mitogenic for connective tissue cells. Approximately 30% of this biological activity can be neutralized by antibodies specific for human PDGF (DiCorleto, 1984).

PDGF has been described as a mitoattractant due to its chemotactic effect on connective tissue cells at lower concentrations (0.1–1.0 nM) and its mitogenic effect on these cells at higher concentrations (0.5–5 nM) (Grotendorst and Martin, 1986). Because of the dual biological activity of this molecule, PDGF is believed to be a major factor involved in the normal healing of wounds and pathologically contributing to the lesions of atherosclerosis, fibrotic diseases, and oncogenesis. PDGF was originally identified (Ross et al., 1974; Kohler and Lipton, 1974) and purified (Antoniades et al., 1979; Heldin et al., 1981) from the alpha-granules of human platelets. Platelet PDGF is a dimeric molecule that migrates on SDS-polyacrylamide gels at ~30 kD. Reduction of interchain disulfide bonds yields A chain (17 kD) and B chain (14 kD) monomers that are not biologically active (Antoniades, 1981; Antoniades et al, 1979; Grotendorst et al., 1982). Conditioned media from cultures of human umbilical vein endothelial (HUVE) cells contain factors that compete with platelet PDGF for binding to the PDGF cell surface receptor of fibroblasts and demonstrate PDGF-related biological activity (DiCorleto, 1984). HUVE cells express both the A and B chain genes of PDGF (Collins et al., 1985; Collins et al., 1987). Since both PDGF gene transcripts are present, the exact nature of the proteins responsible for the PDGF-related mitoattractant activity found in HUVE cell-conditioned media was uncertain. The secreted molecules could be AB heterodimers or AA or BB homodimers. All three isoforms have been purified from natural sources and are biologically active (Heldin et al. 1986; Betsholtz et al. 1986; Stoobant and Waterfield, 1984). Our initial interest was to determine the type of PDGF molecules secreted by vascular endothelial cells. During the course of these studies we have identified a new peptide that appears to be responsible for the PDGF-related mitoattractant activity present in the endothelial cell-conditioned media. We have termed this protein connective tissue growth factor (CTGF).[1]

Gary R. Grotendorst's present address is Department of Cell Biology and Anatomy, University of Miami School of Medicine, 1600 N.W. 10th Ave., Miami, FL 33101.

1. *Abbreviations used in this paper*: CTGF, connective tissue growth factor; HUVE, human umbilical vein endothelial.

Materials and Methods

Cells

HUVE cells were isolated from fresh human umbilical cords by collagenase perfusion (Jaffe, 1987) and maintained in medium 199 with 20% FCS, 0.68 mM L-glutamine, 20 μg/ml Gentamicin, 90 μg/ml porcine heparin (Sigma Chemical Co., St. Louis, MO), and 50 μg/ml endothelial cell growth supplement (Sigma Chemical Co.). Cells used for media collection were third passage cells. Cells were identified as endothelial cells by their nonoverlapping cobblestone morphology and by positive staining for factor-VIII related antigen. NRK cells were obtained from American Type Culture Collection (Rockville, MD), NIH/3T3 cells were a gift from S. Aaronson (National Cancer Institute, Bethesda, MD), and both cell lines were maintained in DMEM, 10% FCS, 20 μg/ml Gentamicin. Fetal bovine aortic smooth muscle cells were obtained from tissue explants as previously described (Grotendorst et al., 1981) and maintained in DMEM, 10% FCS, 20 μg/ml Gentamicin, and used in assays at second or third passage.

Growth Factors and Antibodies

Human PDGF was purified to homogeneity from platelets as described previously (Grotendorst, 1984). Recombinant AA, BB, and AB chain dimeric PDGF molecules were obtained from Creative Biomolecules (Hopkinton, MA). FGF was obtained from Sigma Chemical Co. Purified PDGF or synthetic peptides containing the amino and carboxyl sequences of the mature PDGF A and B chain molecules were used to raise antibodies in goats. Goats were immunized with 20 μg of purified PDGF or 50 μg of synthetic peptide in Freunds complete adjuvant by multiple intradermal injections. Immune sera were collected 7 d after the fourth rechallenge (in Freunds incomplete adjuvant) and subsequent rechallenges. The anti-PDGF antibody did not show any cross-reactivity to TGF-β, EGF, or FGF in immunoblot analysis. The antipeptide antibodies were sequence specific and did not cross-react with other synthetic peptide sequences or with recombinant PDGF peptides that did not contain the specific antigenic sequence. This was determined by Western blot and dot blot analysis.

Antibody Affinity Column

Goat anti-human PDGF IgG (150 mg) was covalently bound to 25 ml of Affi-Gel 10 support (Bio-Rad Laboratories, Cambridge, MA) according to the manufacturer's instructions with a final concentration of 6 mg IgG/ml gel. The column was incubated with agitation at 4°C for 18 h with 1 liter of HUVE cell media that had been conditioned for 48 h. The gel was then poured into a column (5 × 1.5 cm), washed with 4 vol of 0.1 N acetic acid made pH 7.5 with ammonium acetate, and the antibody-bound PDGF immunoreactive proteins were eluted with 1 N acetic acid. Peak fractions were determined by biological assays and immunoblotting and the fractions pooled.

Biological Assays

Chemotactic activity was determined in the Boyden chamber chemotaxis assay with NIH 3T3 or bovine aortic smooth muscle (BASM) cells as previously described (Grotendorst et al., 1981, 1987). Mitogenic assays were performed using 96-well plates and NRK fibroblasts or NIH 3T3 cells as target cells. The cells are plated in DMEM, 10% FCS, and the NRK cell cultures used 10–14 d after confluence and 3T3 cells made quiescent by incubating for 2 d in serum-free DMEM, 0.2 mg/ml BSA before use. Sample proteins and dilutions of known standards were added to the wells and the plates incubated at 37°C in 10% CO_2, 90% air for 18 h, after which ^{3}H-thymidine at a final concentration of 5 μCi/ml was added and incubated for an additional 2 h. The media was removed, the cells washed, and DNA synthesis determined from the ^{3}H-thymidine incorporation into TCA-precipitable material by scintillation counting.

Gel Electrophoresis and Immunoblotting

Electrophoresis was performed on 12% polyacrylamide gels containing SDS (Laemmli, 1970) unless otherwise stated. Immunoblotting was performed by electroblotting the proteins to a nitrocellulose membrane and incubating the membrane in 50 mM Tris-HCl, pH 7.4, 100 mM NaCl (TBS) with 2% nonfat dry milk at 25°C for 1 h to block nonspecific antibody binding. The blocking solution was removed and the antibody (15 μg/ml) added in TBS containing 0.5% nonfat dry milk and 1 μg/ml sodium azide and incubated overnight at 25°C. The membranes were then washed five times in TBS, 0.5% milk for 10 min each wash and then incubated with alkaline phosphatase–conjugated affinity purified rabbit anti-goat IgG (KPL, Gaithersburg, MA) at a 1:1,000 dilution in TBS containing 0.5% milk at 25°C for 1 h. The filters were then washed with TBS five times, 10 min each time, and the blot was developed using an alkaline phosphatase substrate solution (0.1 M Tris-HCl, pH 9, 0.25 mg/ml nitro blue tetrazolium, 0.5 mg/ml 5-bromo-4-chloro-3-indolyl phosphate).

Receptor Competition Assays

Assays were performed using confluent cultures of NIH 3T3 cells in 24-well plates (Costar Data Packaging Corp., Cambridge, MA) grown in DMEM, 10% FSC, 10 μg/ml Gentamicin. The growth media was removed and the cells were washed twice with serum-free DMEM, 0.2 mg/ml BSA, and the plates placed on ice for 30 min in serum-free DMEM, 0.2 mg/ml BSA. Test samples and controls were made up in serum-free DMEM, 0.2 mg/ml BSA containing 5–10 ng/ml of HUVE affinity-purified proteins and a serial dilution of one of the recombinant PDGF isoforms in a concentration range of 300–16 ng/ml. 1-ml aliquots of the samples were placed into wells of the 24-well plates and incubated on ice on a platform rocker for 2 h. After the incubation period, the cells were washed three times for 10 min each on ice with PBS. The proteins bound to the surface of the cells were eluted with 500 μl of 1 N acetic acid for 10 min. The acetic acid elution samples were lyophilized, resuspended in 5 mM HCL, run on 12% polyacrylamide gels, and immunoblotted to nitrocellulose using the anti-PDGF antibody.

RNA Isolation and Northern Blotting

Total RNA was isolated from cells in monolayer culture cells by the method of Chomczynski and Sacchi (1987). Lyophilized RNA was resuspended in gel loading buffer containing 50% formamide and heated at 95°C for 2 min before loading (20 μg per lane total RNA) onto 2.2 M formaldehyde, 1% agarose gels and run at 50 V. Integrity of RNA was determined by ethidium bromide staining and visualization of 18S and 28S rRNA bands. After electrophoresis the RNA was transferred to nitrocellulose by blotting overnight with 10× SSC. The nitrocellulose was air dried and baked at 80°C for 2 h in a vacuum oven. Hybridization was performed overnight at 46°C with the addition of 5 × 10^5 CPM per ml of ^{32}P-labeled probe. Normally, for Northern blots the entire plasmid was labeled and used as a probe. Labeling was done with a random primer labeling kit from Boehringer Mannheim Biochemicals (Indianapolis, IN) according to instructions provided. After hybridization, membranes were washed twice in 2× SSC, 0.1% SDS for 15 min each at room temperature, once for 15 min in 0.1× SSC, 0.1% SDS room temperature, and a final 15-min wash in 0.1× SSC, 0.1% SDS at 46°C. Blots were autoradiographed at −70°C on Kodak X-omat film.

Library Screening, Cloning, and Sequencing

Standard molecular biology techniques were used to subclone and purify the various DNA clones (Sambrook et al., 1989). Clone DB60 was picked from a lambda-gt11 HUVE cell cDNA library by induction of the fusion proteins and screening with anti-PDGF antibody. Plaques picked were rescreened and positive clones replated at low titer and isolated.

The EcoRI insert from clone DB60 was cloned into the M13 phage vector and single-stranded DNA obtained for clones with the insert in opposite orientations. These M13 clones were then sequenced by the dideoxy method (Sanger, 1977) using the Sequenase kit (United States Biochemical Corp., Cleveland, OH) and ^{35}S-dATP (duPont Co., Wilmington, DE). Both strands of DNA for this clone were completely sequenced using primer extension and both GTP and ITP chemistry. The sequencing reactions were done according to the manufacturer's instructions. Aliquots of the sequencing reactions were run on both 6% acrylamide (16 h) and 8% acrylamide (6 h) gels, vacuum dried, and autoradiographed for at least 18 h.

The cDNA fragment from clone DB60 was ^{32}P-CTP labeled and used to rescreen the HUVE cell cDNA lambda gt11 library. Several clones were picked and the largest, the 2,100-bp clone designated DB60R32, was subcloned into Bluescript phagemid. Subclones were made of PstI, KpnI, and EcoRI/KpnI restriction fragments, also in Bluescript. These subclones were sequenced by double-stranded plasmid DNA sequencing techniques using Sequenase as described above. The 1,458-bp EcoRI/KpnI clone containing the open reading frame was subcloned into M13 mp18 and M13 mp19, and both strands of DNA were completely sequenced using single-stranded DNA sequencing techniques with primer extension and both GTP and ITP chemistry.

In Vitro Transcription and Translation

In vitro transcription reactions were done using the 2,100-bp cDNA clone DB60/32 in the Bluescript KS vector. The plasmid was cut with XhoI, which cuts the plasmid once in the multiple cloning site of the vector 3' to the cDNA insert. The T7 promoter site located 5' to the cDNA insert was used for transcription. The in vitro transcriptions were done with a kit supplied with the Bluescript vector by Stratagene Cloning Systems and the manufacturer's instructions were followed.

In vitro translation reactions were done using nuclease-treated rabbit reticulocyte lysate and ^{35}S-cysteine in a cysteine-free amino acid mix for labeling of the peptide. The reactions were done with a kit supplied from Promega Biotec (Madison, WI) and the manufacturer's instructions were followed. The reactions were done in a final volume of 50 μl containing ^{35}S-cysteine 1 mCi/ml (1,200 Ci/mMole, DuPont Co.), and serial dilutions of mRNA from the in vitro transcription reactions in concentrations ranging from 50 to 500 ng per reaction tube. The reactions were incubated at 30°C for 60 min. Aliquots of the reactions were run reduced or nonreduced on 12% polyacrylamide electrophoresis gels, dried, and autoradiographed.

Bacterial expression of immunoreactive CTGF peptide was accomplished by subcloning clone DB60R32 into the EcoRI site of the pET 5 expression vector (Studier et al., 1990) in both sense and inverse orientations (as determined by restriction enzyme digest analysis). Cultures of cells were grown in M9 media to an OD 600 of 0.7 and the media made 0.4 mM IPTG and incubation continued for 2 h. The cells were pelleted, lysed, inclusion bodies removed by centrifugation, and aliquots of the pellet extracts run on 12% polyacrylamide gels and immunoblotted using the anti-PDGF antibody.

For expression in *Xenopus* oocytes, mature *X. laevis* females were obtained from Nasco (Fort Atkinson, WI) and maintained at room temperature. Frogs were anesthetized by hypothermia and the ovarian tissue was surgically removed. Ovarian tissue was minced and digested the 0.2% collagenase (type II; Sigma Chemical Co.) in OR-2 without calcium (Wallace et al. 1973) for 2-3 h. Unblemished stage VI oocytes (Dumont, 1972), 1.3-mm diameter, were then carefully selected and microinjected.

Stage VI oocytes (5–10 at a time) were placed on a hollowed plexiglass platform and drained of excess OR-2 solution. Approximately 50 nl of sample containing 10 ng of RNA was injected into the animal pole just above the oocyte equator using a Leitz system microinjector. After injection, oocytes were returned to OR-2 buffer with 0.1% BSA and incubated for 24 h at 25°C. Viable oocytes were then pooled and extracted by homogenization in 100-mm NaCl, 10 mm Tris, pH 7.5, with ter strokes of a Dounce homogenizer (20 μl/oocyte). The homogenate was then mixed with an equal volume of freon to remove pigment and lipid and centrifuged at 10,000 rpm for 30 s to separate the phases. The top aqueous phase was removed and tested for chemotactic activity using NIH 3T3 cells as described above.

Results

Identification and Partial Purification of PDGF-immunorelated Mitogen from HUVE Cells

Initial studies of the PDGF-related growth factors secreted by HUVE cells were done by removing the serum-containing growth media from confluent cultures of cells and replacing it with serum-free media. Aliquots of this media were removed periodically and the proteins immunoblotted using an antibody specific for human platelet PDGF (Fig. 1). This antibody does not cross-react with any other known growth factors and is able to detect <500 pg of dimeric PDGF or 10 ng of reduced, monomeric A or B chain peptide on immunoblots. The results indicated constitutive secretion of several species of molecules which are immunologically similar to platelet PDGF but are of higher relative molecular mass (36–39 kD) than the expected 30–32 kD molecular mass of platelet PDGF or A chain or B chain homodimers. Chemotactic and mitogenic assays performed with this serum-free conditioned media indicated the total biological activity present was equivalent to 15 ng/ml of platelet PDGF aft r a 48-h conditioning period (Fig. 2). Incubation of the media

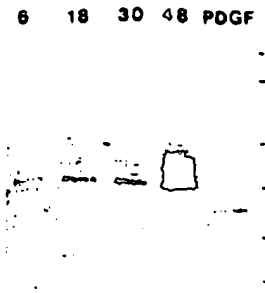


Figure 1. Constitutive secretion of PDGF-immunoreactive factors by HUVE cells. HUVE cells were grown to confluence in 6-well plates. The growth media was removed, cells were washed with PBS, and 1 ml of serum-free media was added to each well. The media was removed after conditioning for the period of time indicated (hours), dialyzed against 1 N acetic acid, and lyophilized. The samples were then run on 12% PAGE, electroblotted to nitrocellulose, and visualized with the antihuman PDGF antibody. 5 ng of purified platelet PDGF was run as reference. Positions of mol wt markers (Bio-Rad Laboratories) are indicated at right.

with 30 μg/ml of anti-human PDGF IgG neutralized ~20–30% of the mitogenic activity and similar amount of the chemotactic activity. This is in agreement with previous reports (DiCorleto, 1984).

The presence in HUVE culture media of several species of PDGF-immunoreactive molecules was unexpected, particularly molecules of higher molecular weight than those of the A and B chain dimeric molecules anticipated to be produced and secreted by endothelial cells (Collins et al., 1987; Sitaras et al., 1987). To obtain greater amounts of the PDGF-like proteins for further analysis, the HUVE cells had to be kept in media containing 20% FCS, as the cells begin to die after 24 h in serum-free or low serum media. The PDGF-immunoreactive proteins were partially purified from the serum-containing media by use of an antibody affinity column made with the anti-human PDGF IgG and an Affi-Gel 10 support (BioRad Laboratories). When aliquots of the partially purified proteins were assayed for chemotactic and mitogenic activity, all biological activity could be neutralized by prior incubation of the proteins with the antihuman PDGF antibody (Fig. 2). This indicated that the only biologically active molecules present in the partially purified media proteins were PDGF-immunorelated molecules. Aliquots of the partially purified proteins were immunoblotted using the same anti-PDGF antibody and the data indicated the presence of the higher molecular weight molecules observed in the serum-free conditioned media (Fig. 3). The major species secreted migrates on polyacrylamide gels at 36 kD and comprises at least 50% of the total immunoreactive protein purified from conditioned media. The immunoreactive species migrating at 37 and 39 kD constitute most of the remaining immunoreactive protein. A similar pattern is seen with proteins labeled with ^{35}S-Cysteine and affinity purified with the anti-PDGF IgG immunoaffinity column (data not shown). Less than 15% of the total affinity-purified proteins comigrate with purified platelet PDGF or recombinant PDGF isoforms. Prior incubation of the antibody with purified PDGF (300 ng PDGF/2 μg IgG) blocked antibody binding to all f the molecules, indicating shared antigenic determinants with dimeric platelet PDGF (Fig. 3, lane *4*). Interestingly, when the antibody was blocked with recombinant AA, BB, or AB dimers, antibody binding to the HUVE-secreted proteins was inhibited equally by all three

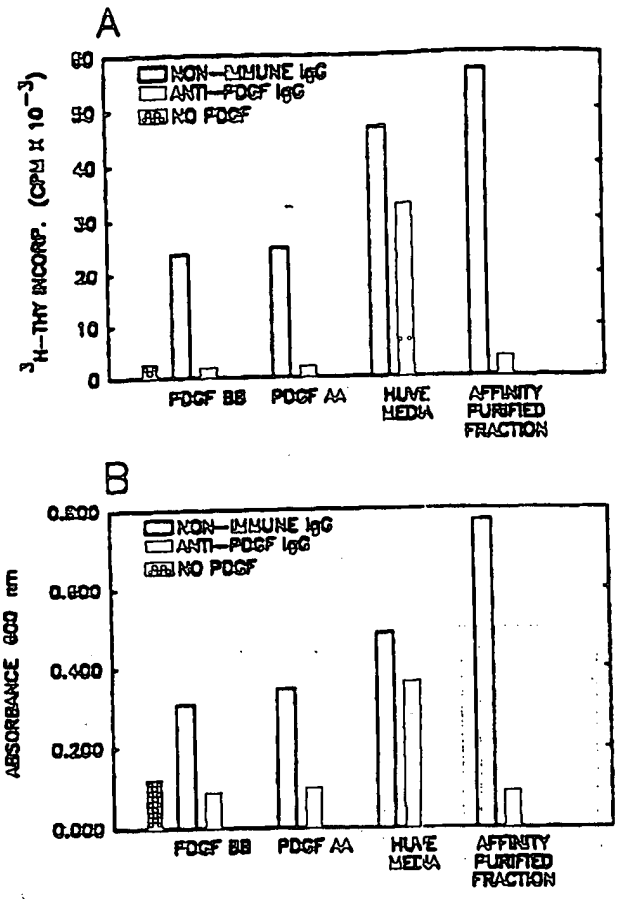


Figure 2. Chemotactic and mitogenic assays of HUVE cell-conditioned media and affinity-purified PDGF immunoreactive factors. (*A*) Mitogenic assay performed as described using NRK cells as target cells. PDGF BB is 5 ng/ml. PDGF AA is 10 ng/ml. HUVE media is 250 μl of HUVE cell serum-free conditioned media (48 h) which was dialyzed against 1 N acetic acid, lyophilized, and resuspended in DMEM before addition to test wells. Affinity-purified fraction is 5 μl/ml of combined, concentrated major pool from Affi-Gel 10 affinity column. Anti-PDGF IgG or nonimmune IgG (30 μg/ml) was added to the samples and incubated 18 h at 4°C before testing in the mitogenic assay. Data points in *A* and *B* represent the mean of triplicate samples and the standard deviation is <5%. The experiments were repeated at least three times with similar results. (*B*) Chemotactic assays were performed as described under Materials and Methods using NIH 3T3 cells as target cells. PDGF BB is 5 ng/ml. PDGF AA is 10 ng/ml. HUVE media is serum-free DMEM, 0.2 mg/ml conditioned for 48 h. Affinity-purified fraction is 2.5 μl/ml of combined, concentrated major pool from Affi-Gel affinity column. Antibody neutralization is performed as described under *A*.

dimeric forms, suggesting that the antibody recognizes common epitopes present on all three PDGF dimers and the HUVE-secreted molecules (our unpublished observations). To insure that none of the antibody binding molecules detected on Western blots were derived from FCS or other additives in the culture media, a new, unused antibody affinity column was made and media not conditioned by cells was processed exactly as the conditioned media. No PDGF-immunoreactive molecules were detected in the fractions from this column by immunoblot (Fig. 3, lane 5) and no

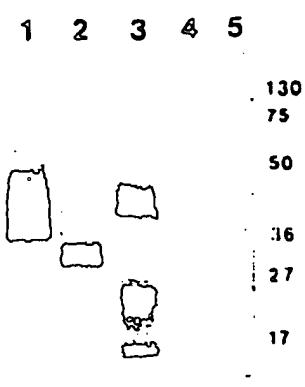


Figure 3. Immunoblot of HUVE cell-secreted PDGF-related factors. Proteins were run on a 12% polyacrylamide gel electroblotted to a nitrocellulose membrane and the immunoreactive factors visualized with antihuman PDGF IgG as described under Materials and Methods. (lane *1*) 5 μl of affinity-purified HUVE cell-secreted proteins. (lane *2*) 10 ng of synthetic AA homodimer (*top band*) and 10 ng of BB homodimer (*lower band*). (lane *3*) 50 μl of reduced HUVE affinity-purified proteins. (lane *4*) 5 μl of HUVE affinity-purified HUVE proteins but with antihuman PDGF antibody blocked with 300 ng of PDGF. (lane *5*) Control column of 10 μl of anti-PDGF Affi-Gel 10 affinity-purified protein fraction from media which was not conditioned by cells.

biological activity was detected (data not shown). When platelet PDGF or the recombinant dimers are reduced with 100 mM DTT, monomeric A chain (17 kD) and B chain (14 kD) peptides are observed on immunoblots. Treating the HUVE molecules in a 100-mM DTT sample buffer results in slower migration of the major immunoreactive peptides on polyacrylamide gels (Fig. 3, lane *3*). Most of the immunoreactive molecules migrate at 38–39 kD and less intense bands are observed at 25 and 14 kD. It is necessary to run at least 10 times as much reduced protein as nonreduced in order to detect the reduced molecules. This is consistent with the affinity of our antibody for monomeric forms of the PDGF A and B chain peptides. These data indicated that the major species in the PDGF-related affinity-purified proteins from conditioned media of HUVE cells was monomeric peptide which migrates on acrylamide gels at an apparent molecular mass of 36 kD nonreduced and 38 kD when reduced.

Major Chemotactic and Mitogenic Activity Is Produced by 36-kD Peptide and Not PDGF Peptides

To determine if the chemotactic and mitogenic activities observed in the partially purified media proteins were from molecules containing the PDGF A and B chain peptides or were the products of molecules that do not contain these sequences, biological assays were performed with serial dilutions of the affinity-purified media proteins and serial dilutions of recombinant PDGF AA and BB homodimers and the AB heterodimer. (Fig. 4). Sufficient quantities of the samples were prepared to perform the mitogenic and chemotactic assays and the immunoblots with aliquots of each dilution sample. The mitogenic activity of the HUVE affinity-purified factors observed was comparable to the activity elicited by all three recombinant PDGF dimers. The chemotactic activity was comparable to the AB heterodimer, producing less response than the BB homodimer and greater response than the AA homodimer. When the biological activity of the samples was compared with immunoblots of equivalent amounts of the same samples, no A chain nor B chain molecules were detected in the test samples (Fig. 4). These data demonstrate the major biological activity present in the anti-PDGF affin-

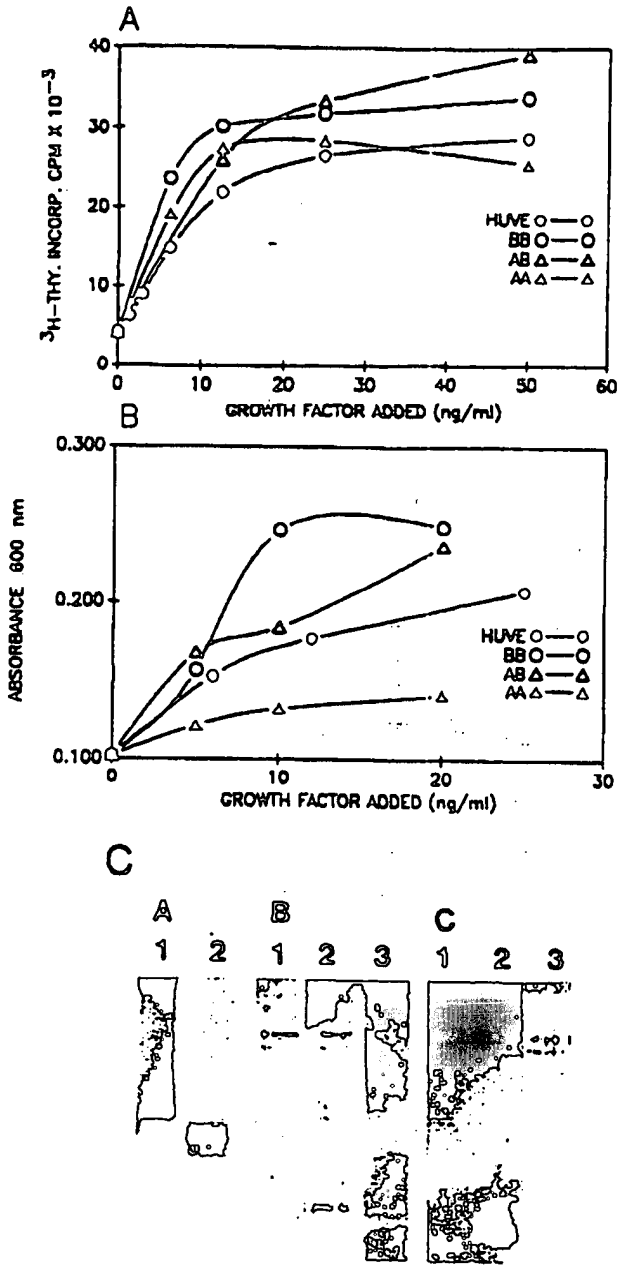


Figure 4. Biological assays and immunoblots of serial dilutions of HUVE cell affinity-purified media proteins and recombinant PDGF standards. (*A* and *B*) Filled circle is HUVE cell affinity-purified media proteins; open circle is BB homodimer; open triangle is AB heterodimer; closed triangle is AA homodimer. Mitogenic assay (*A*) and chemotactic assay (*B*) were performed as described with NIH 3T3 cells. The test samples in both biological assays and the immunoblot are equal aliquots of the same dilution sample. Data points represent the mean of triplicate samples with SD <10%. (*C*) (*immunoblot A*) Primary antibody is anti-human PDGF IgG. (lane *1*) 20 ng HUVE purified media proteins. (lane *2*) 20 ng AB heterodimer. (*immunoblot B*) Primary antibody is antiamino terminal A chain serum. (lane *1*) 20 ng of HUVE-purified media proteins, reduced. (lane *2*) 20 ng AA homodimer, reduced. (lane *3*) 1.25 ng AA homodimer, reduced. (*immunoblot C*) Primary antibody is anticarboxy terminal B chain serum. (lane *1*) 20 ng HUVE-purified media proteins, reduced. (lane *2*) 20 ng BB homodimer, reduced. (lane *3*) 2.5 ng BB homodimer, reduced.

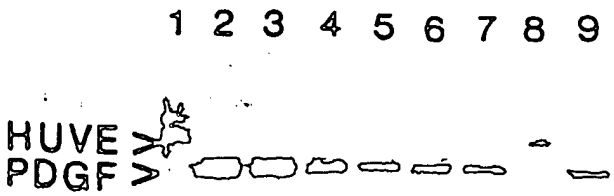


Figure 5. PDGF cell surface receptor-binding competition assay with NIH 3T3 cells and HUVE affinity-purified proteins competing with recombinant PDGF BB homodimer. Lane 1 contains 10 ng of HUVE affinity-purified proteins and lane 9 contains 1.0 ng of recombinant PDGF BB homodimer. Lanes 2–8 are proteins dissociated in one well of a 24-well plate from the cell surface of NIH 3T3 cells with acetic acid. The cells in each well were incubated for 2 h at 4°C with serum-free DMEM containing 10 ng of affinity-purified protein from HUVE cell-conditioned media and varying concentrations of recombinant PDGF BB. The concentration of PDGF in lane 2 is 300 ng; lane 3, 150 ng; lane 4, 75 ng; lane 5, 37.5 ng; lane 6, 18.75 ng; lane 7, 9.4 ng; and lane 8 contains no addition of PDGF.

ity-purified fraction cannot be accounted for by PDGF A or B chain–containing molecules and imply that the major PDGF-immunoreactive protein species present in these samples (the 36-kD peptide) is biologically active and does not contain amino acid sequences found in the amino and carboxy terminals of the PDGF A or B chain peptides.

To substantiate the binding of the endothelial cell molecules to the PDGF cell surface receptors, competitive receptor binding assays were performed. Because immunoblots of the affinity-purified HUVE cell–secreted proteins indicated the presence of multiple PDGF immunoreactive molecules, ^{125}I-labeled PDGF competition assays could not be used since this would not indicate which molecules in this mixture were competing for binding of the labeled PDGF for the receptors on the target cells. Since the isoforms of PDGF and the major PDGF immunorelated protein secreted by HUVE cells are of different molecular weights, we were able to demonstrate receptor binding competition on immunoblots. Direct binding of the anti-PDGF immunoreactive peptides to NIH 3T3 cells was demonstrated by incubating monolayers of the 3T3 fibroblasts with the anti-PDGF affinity-purified proteins (10 ng/ml) for 2 h at 4°C. Bound peptides were released by washing of the cell layer with 1 N acetic acid and quantitated by immunoblot analysis using anti-PDGF IgG (Fig. 5). The data show that the 36-kD immunoreactive peptide binds to cell surface of NIH 3T3 cells. This binding can be competed by increasing concentrations of recombinant PDGF BB added to the binding media. These data suggest that the CTGF peptide binds to specific cell surface receptors on NIH 3T3 cells and that PDGF BB can compete with this binding. Whether CTGF binds to a certain class of PDGF receptors or whether there is some cross reactivity of PDGF BB with CTGF receptors that are distinct from PDGF receptors is not clear from these results and will require a more in-depth study.

Cloning Expression and Sequencing of the cDNA for CTGF

To further characterize these PDGF-related molecules, we

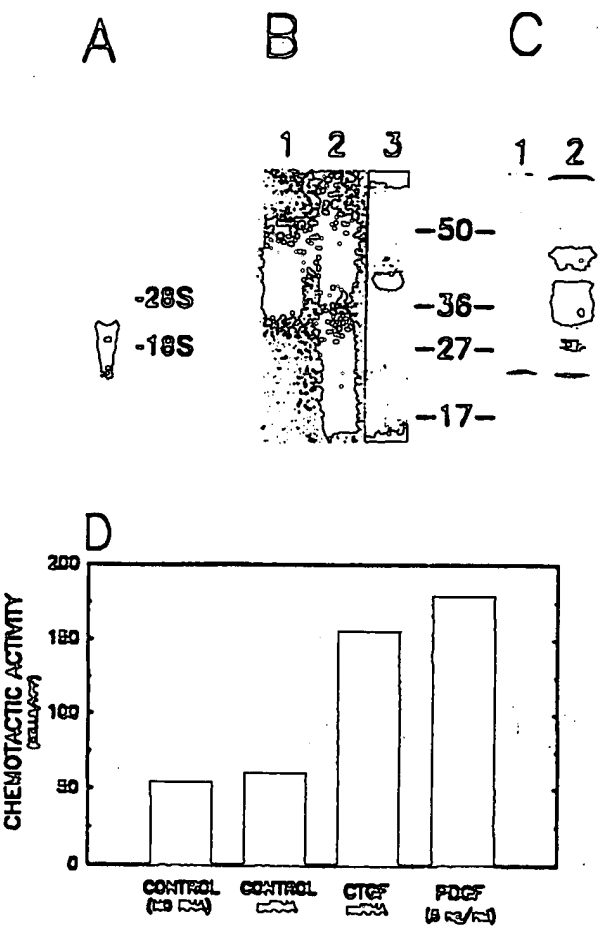


Figure 6. Northern blot of CTGF transcript in HUVE total RNA and in vitro translation product of cDNA clone DB60R32 and immunoreactive peptide from prokaryotic expression. (*A*) 20 μg of total RNA from cultured HUVE cells was run on a 1.5% agarose gel, transferred to nitrocellulose, and probed with ^{32}P-labeled DB60 clone. Autoradiograph indicates hybridization to a 2.4-kb transcript. (*B*) Lanes 1 and 2 are an immunoblot using the antihuman PDGF antibody with (*1*) 10 ng nonreduced and (*2*) 100 ng reduced of HUVE cell affinity-purified proteins from conditioned media. Lane 3 is an autoradiograph of 4 μl of a 50-μl rabbit reticulocyte in vitro translation reaction which incorporated 1 μg of RNA transcript from an in vitro transcription reaction using the 2,100-bp clone DB60R32 in Bluescript phagemid. Autoradiograph was exposed for 24 h. (*C*) Clone DB60R32 was cloned into the EcoRI site of pET5 prokaryotic expression vector in both sense and inverse orientations. Protein expression was induced by IPTG for 2 h and 20-μl aliquots of cell pellet extract were run on 12% polyacrylamide gels and immunoblotted using the anti-PDGF antibody. Lane 1 is the antisense control and lane 2 is the sense strand peptide produced by clone DB60R32. (*D*) Capped mRNA was prepared by in vitro transcription of the DB60R32 clone in Bluescript phagemid after restriction with XhoI. Oocytes were microinjected with buffer (control, no RNA) or 10 μg of either control or CTGF mRNA. Total oocyte protein was extracted after 24 h. Chemotaxis assays were performed as described under Materials and Methods with NIH 3T3 cells using 25 μg of oocyte total protein in each sample. PDGF BB (5 ng/ml) positive control is shown for comparison. The results represent the average of triplicate samples with a variation of <10%. These studies were repeated twice.

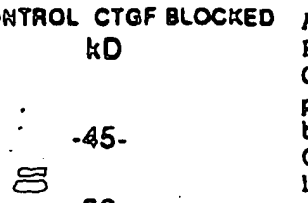


Figure 7. Blocking of anti-PDGF reactivity of native CTGF by recombinant CTGF produced in *E. coli*. Western blots were performed using CTGF isolated from endothelial cell-conditioned media and anti-PDGF IgG as described under Materials and Methods. Identical samples of CTGF were probed with anti-PDGF IgG that had been incubated with control bacterial cell extract (non-CTGF producing) (non-blocked) or with an equal amount (50 μg total protein) of extract from a recombinant strain producing CTGF (blocked). These are the same recombinant pET-CTGF and control bacterial cells used in the experiment described in Fig. 6 *C*. The recombinant CTGF completely blocked the immunoreactivity of the anti-PDGF with the native CTGF isolated from the conditioned media.

first attempted to obtain sufficient quantities of the CTGF protein for amino acid sequencing. However, the low concentrations of CTGF in the conditioned media of HUVE cell cultures and the costly and time consuming techniques involved in obtaining and culturing these cells made protein purification to homogeneity and amino acid sequencing impractical. Therefore, we used the anti-PDGF antibody to screen an HUVE cell cDNA library made in the expression vector lambda gt11 (a gift from T. Collins, Harvard). Over 500,000 recombinant clones were screened. Several clones which gave strong signals with the anti-PDGF antibody in the screening process were purified and subcloned into the M13 phage vector and partial sequence data obtained by single stranded DNA sequencing. A search of the GenBank DNA sequence data base indicated that two of the clones picked contained fragments of the PDGF B chain cDNA open reading frame sequence. One of these clones was similar to a 1.8-kb insert previously isolated by Collins et al. (1985) using a c-sis cDNA probe. A third clone of 500 bp was completely sequenced and no match was found in a homology search of all nucleotide and amino acid sequences in GenBank (CEF 10 sequence was not available at that time). This clone was designated DB60. Anti-PDGF antibody binding to the fusion protein produced by the clone DB60 was completely blocked by the affinity-purified proteins (not shown). A ^{32}P-labeled probe was made of DB60 and used on a Northern blot of 20 μg of total RNA isolated from HUVE cells (Fig. 6 *A*). The blot indicated probe hybridization with an mRNA of 2.4 kb, which is a message of sufficient size to produce the proteins in the 38-kD molecular mass range seen on the immunoblots of the affinity-purified proteins. The DB60 clone was used to rescreen the HUVE cell cDNA lambda gt11 library and the largest clone isolated contained a 2,100-bp insert designated DB60R32. A probe made with the 2,100-bp EcoRI insert of clone DB60R32 also hybridized with a single 2.4-kb message in a Northern blot of total RNA from HUVE cells (not shown). To determine the size of the peptide encoded in the open reading frame of DB60R32, the 2,100-bp insert was cloned into the Bluescript vector and transcribed in vitro using T7 polymerase and the mRNA transcript translated in vitro using a rabbit

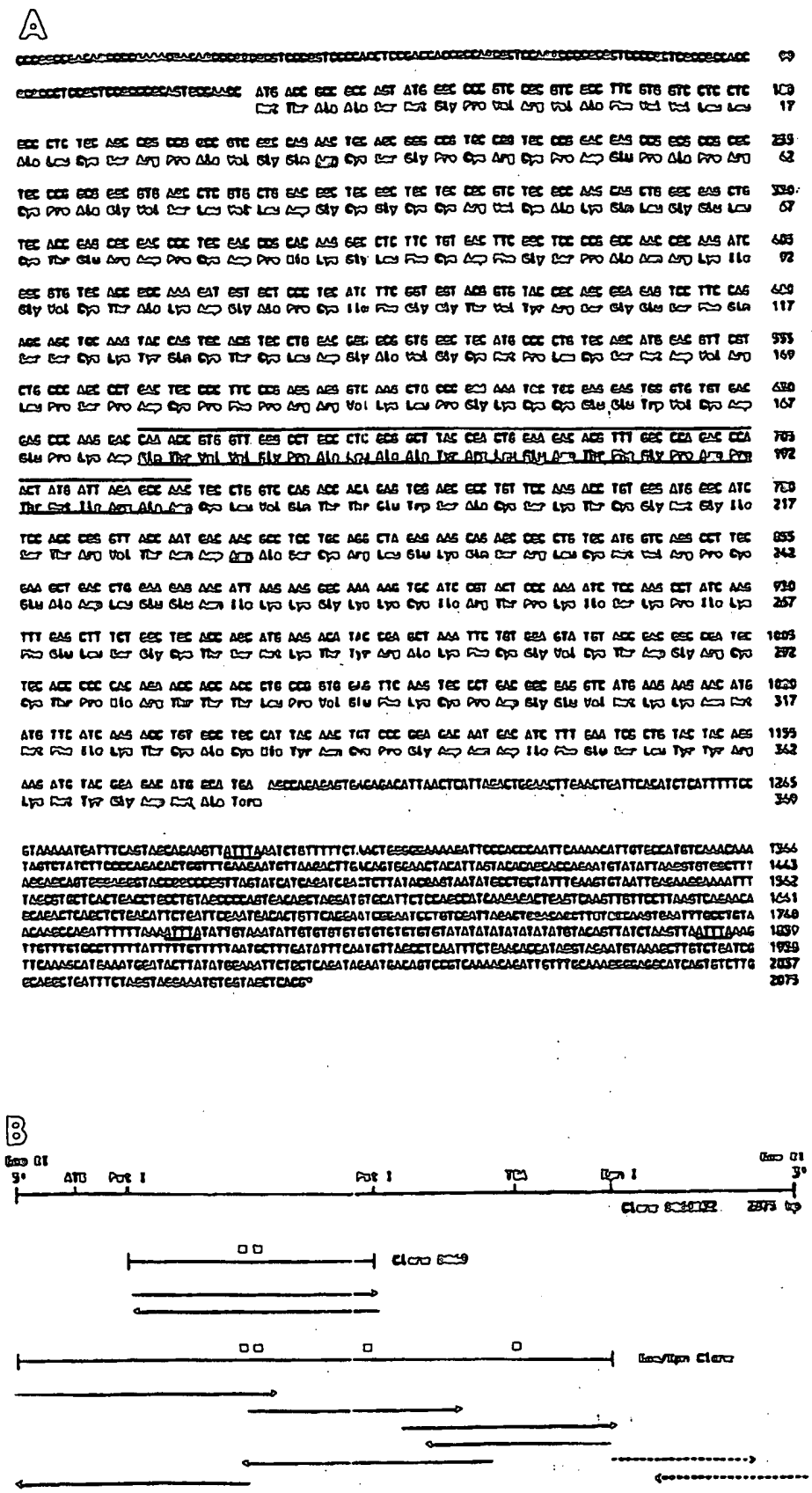


Figure 8. (*A*) Nucleotide sequence and amino acid translation of CTGF cDNA from human vascular endothelial cells. The open reading frame for the CTGF protein extends from the ATG initiation site at nucleotide 130 to the TGA site at 1177. The possible glycosylation sites at asparagines 28 and 225 are underlined. The putative alternative splicing region is boxed. The 3' region contains three ATTTA sites which are underlined. (*B*) Sequencing map of CTGF cDNA clone. The top line represents the 2,075-bp DB60R32 clone containing the open reading frame of CTGF. Restriction enzyme sites used in subcloning are indicated as are the ATG initiation site and the TGA termination codon. Squares indicate locations of primers used in sequencing. Solid arrows indicate single-stranded DNA sequencing data and broken arrows indicate area in 3' regi n sequenced by double-stranded DNA methods.

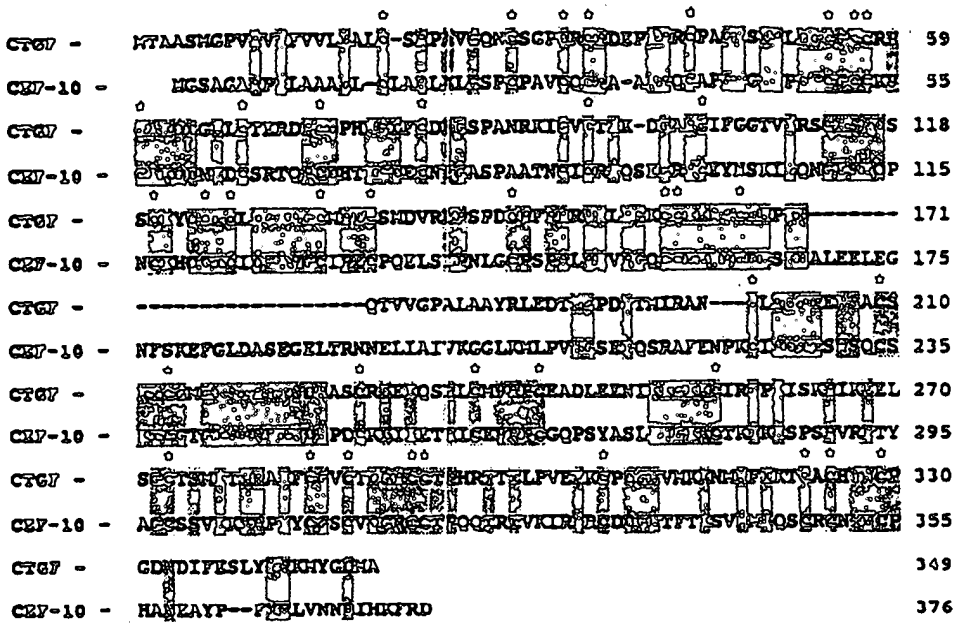


Figure 9. Comparison of amino acid sequences of the translated cDNA for CTGF and the CEF-10 mRNA translation product. The translated cDNA for human CTGF and avian CEF-10 have a 45% overall homology and a 52% homology if the putative alternative splicing region is deleted. This region is between amino acids 171 (aspartic acid) and 199 (cysteine) in the CTGF sequence. Homologous regions are shaded and the aligned cysteines are marked with an asterisk.

reticulocyte lysate and ^{35}S-cysteine. Aliquots of the translation reaction were reduced with 50 mM DTT and run on 12% polyacrylamide gels. The data indicate a protein product which migrates at an apparent molecular mass of 38 kD and comigrates with the reduced affinity-purified proteins observed on immunoblots (Fig. 6 *B*). The clone DB60R32 was subcloned into the EcoRI site of the prokaryotic expression vector pET 5 in both orientations and transfected into *E. coli* strain HMS174. The CTGF protein was induced by IPTG and aliquots of the cell lysate were immunoblotted (Fig. 6 *C*). The protein produced by clone DB60R32 in the sense orientation produced anti-PDGF immunoreactive peptides in the 36–39 kD molecular mass range, while the antisense control produced no immunoreactive peptides. The recombinant peptides produced in the *E. coli* system completely blocked the anti-PDGF reaction with the CTGF peptides present in conditioned media (Fig. 7).

Bioassays of extracts of the recombinant bacterial cells that were producing the CTGF peptide were negative. We then evaluated several eukaryotic expression systems, including microinjection of *Xenopus* oocytes. Injection of *Xenopus* oocytes with 10 ng of RNA preparations derived by in vitro transcription of the DB60R32 clone resulted in the production of a fibroblast chemotactic activity. (Fig. 6 *D*). Control injected cells did not produce this activity. These results indicate that the open reading frame of the DB60R32 clone encodes a protein with chemotactic activity for fibroblastic cells as does CTGF.

The 2,100-bp insert of clone DB60R32 was sequenced initially by subcloning of Pst I and KpnI restriction fragments into Bluescript and using double-stranded dideoxy methods. This determined an open reading frame of 1,047 base pairs and oriented the DB60 insert to the larger cDNA (Fig. 8). An EcoRI/KpnI fragment containing the entire open reading frame was inserted into M13 mp18 and M13 mp19, and both strands of the DNA were sequenced with single-stranded dideoxy methods by primer extension using both GTP and the GTP analogue ITP. The cDNA nucleotide sequence of the open reading frame encoded a 38,000 mol wt protein, confirming our cell-free translation results and matching the size of the immunopurified peptides (Fig. 8). A new search of the GenBank data base revealed that this cDNA had a 50% nucleotide sequence homology with CEF-10 mRNA, one of the immediate early genes induced in v-src–transformed chicken embryo fibroblasts (Simons et al., 1989) (Fig. 9).

Discussion

The data presented here indicate that the major connective tissue mitoattractant secreted by HUVE cells is a 38-kD monomeric molecule which appears to be antigenically and functionally related to PDGF, but is not a product of the PDGF A or B chain genes. We find that HUVE cells also secrete both PDGF A and B chain molecules, but at much lower concentrations than the 38-kD monomer and these peptides contribute only a minor fraction (<10%) of the total PDGF-related chemotactic and mitogenic activity secreted by these cells under the conditions tested. In the biological assays performed with the affinity-purified proteins from HUVE media, biological activity equivalent to 20 ng/ml of PDGF was obtained at low concentrations of the CTGF molecule where no detectable (<2 ng) PDGF A or B chain molecules were present. All biological activity could be removed by prior incubation of these samples with the anti-PDGF antibody. Because all three isomeric forms of the PDGF dimers blocked antibody binding to the CTGF molecule, it is suggested that there are some common antigenic determinants shared among these proteins although there is little if any peptide sequence homology. The anti-PDGF antibody has high affinity to the nonreduced forms of the PDGF isomers and the CTGF molecule and tenfold less affinity to the reduced forms of these peptides which lack biologically activity. This suggests that there are regions of shared tertiary structure between the PDGF isomers and the CTGF molecule, resulting in common antigenic epitopes and possible receptor binding sites of these molecules.

The sequence of the cDNA for CTGF indicates an open reading frame of 1,047 nucleotides with an initiation site at

position 130 and a TGA termination site at position 1,177 which encodes a peptide of 349 amino acids (Fig. 8). The ATG codon at position 130 and another at position 145 fit the consensus sequence for strong translation initiation sites (Kozak, 1984) and it is assumed that the first ATG is the predominant initiation site. There is a 40% sequence homology between the CTGF cDNA and the cDNA for both the A and B chains of PDGF, suggesting a possible common ancestral gene. The 3' region contains three copies of the pentanucleotide sequence ATTTA shown to be involved in mRNA destabilization and frequently found in cytokine and oncogene mRNA 3' regions (Shaw and Kamen, 1986).

The CTGF open reading frame encodes a peptide that contains 39 cysteine residues, indicating a protein of complex structure with multiple intramolecular disulfide bonds. This may explain the shift to slower mobility observed on polyacrylamide gels after reduction of the molecule with DTT. The amino terminal of the peptide contains a hydrophobic signal sequence indicative of a secreted protein, and there are two N-linked glycosylation sites at asparagine residues 28 and 225 in the amino acid sequence. There is a 45% overall sequence homology between the CTGF peptide and the protein encoded by the CEF-10 mRNA transcript and the homology rises to 52% when a putative alternative splicing region is deleted. All 39 cysteine residues in each peptide can be aligned with few gaps in the sequences (Fig. 9). The region between amino acid residues 171 (aspartic acid) and 199 (cysteine) in the CTGF peptide has no significant homology to the corresponding region in the CEF-10 sequence (amino acids 168–224) and is 28 amino acids shorter. This discrepancy could be due to differences incurred during evolution of the chicken and human genes. However, because this region is bordered by areas of very high homology between the two molecules (>85% identity) it may indicate alternative splicing mechanisms in the expression of this gene. The codons for lysine at residue 170 and aspartic acid at residue 171 together form the AAG/G sequence consistent with a 3' exon/5' exon junction. Alternative splicing is found in other growth factor transcripts such as the PDGF A chain (Collins et al., 1987) and vascular endothelial growth factor (Leung et al., 1989; Keck et al., 1989; Tischer et al., 1989). The biological significance of either of these splicing events has not yet been determined.

Simmons et al. (1989) cloned the CEF-10 mRNA, which was one of 12 identified cDNA sequences transcribed from mRNAs that were induced soon after the production of the src phosphoprotein pp60$^{v\text{-}src}$ in chicken embryo fibroblasts. The CEF-10 mRNA was induced in nontransfected CEF cells by serum. Both the src protein and serum induce the expression of the "immediate early genes" many of which are necessary for the G_0–G_1 transition in the cell cycle (for review see Rollins and Stiles, 1989). One major group of these genes consists of intranuclear DNA-binding proteins including fos, myc, and jun, which are essential for cell cycle regulation. Another major group of genes induced by serum and src encodes secretory proteins with cytokine characteristics such as the JC gene (Cochran et al., 1983), which has a cDNA sequence with significant homology to the cytokines macrophage colony stimulating factor (M-CSF), alpha interferon and interlukin-2 (Rollins et al., 1988), or the KC gene (Cochran et al., 1983) which is homologous to the gro gene in humans (Oquendo et al., 1989; Anisowicz et al., 1987). The gro protein product is related to CEF-4, another of the 12 src and serum-inducible mRNAs cloned from CEF cells (Bedard et al., 1987). We have found that the CTGF gene is rapidly induced by serum in human skin fibroblasts and that cycloheximide treatment does not block this induction (Igarashi, A., and G. R. Grotendorst, manuscript submitted for publication). Thus, the CTGF gene is an immediate early gene which appears to encode a secreted peptide with cytokine activity.

The data presented here suggest that the CTGF molecule has biological activity similar to PDGF and may bind one of the PDGF cell surface receptors. CTGF could function in many biological processes involving the growth of connective tissue. The fact that we find CTGF secreted by vascular endothelial cells indicates that the peptide could be present in serum. The conditions under which the protein would be secreted in vivo are not yet known and we are currently pursuing studies on the regulation of gene expression and protein secretion by cultured endothelial cells. Secretion of CTGF during angiogenesis would facilitate the growth of smooth muscle cells and fibroblasts so that it may play an important role in the control of blood vessel formation during development and wound repair. CTGF could also play a role in atherosclerosis where it could function to recruit smooth muscle cells from the medial layer of the vessel wall into the intima. Other investigators have detected PDGF transcripts in arterial tissues (Barrett and Benditt, 1987), and the PDGF mRNA levels appear to be elevated in atherosclerotic lesions (Barrett and Benditt, 1988). The secretion of CTGF peptides by endothelial cells would also stimulate the chemotaxis and growth of smooth muscle and fibroblasts at the plaque site, thereby aggravating vessel blockage. Whether CTGF production is required for transformation by the src oncogene remains to be determined, but it is interesting that CTGF could function as an autocrine growth factor for src transformed fibroblasts. Direct experiments to determine the transforming potential of this cDNA are currently underway. We are also in the process of constructing vectors for eukaryotic expression systems in order to obtain recombinant CTGF for further analysis. Future experiments using CTGF-specific antibodies and nucleotide probes should help to determine the role of CTGF during the normal biological and pathological processes which involve connective tissue formation.

We gratefully acknowledge the assistance of Linda Gibeaut in the preparation of this manuscript. We also thank Marc Charette of Creative Biomolecules (Hopkinton, MA) for the generous gifts of recombinant PDGFs and F. W. Studier for the pET expression vectors and host bacteria and Tucker Collins (Boston, MA) for the human endothelial cell gt 11 library. Douglass M. Bradham is in the Molecular and Cellular Biology Program at the Medical University of South Carolina, Charleston, SC.

This work was supported by grants from the National Institutes of Health and the Florida High Technology Council.

Received for publication 12 January 1991 and in revised form 31 May 1991.

References

Anisowicz, A., L. Bardwell, and R. Sager. 1987. Constitutive overexpression of a growth-regulated gene in transformed Chinese hamster and human cells. *Proc. Natl. Acad. Sci. USA.* 84:7188–7192.

Antoniades, H. N. 1981. Human platelet derived growth factor (PDGF): purification of PDGF I and PDGF II and separation of their reduced subunits.

Proc. Natl. Acad. Sci. USA. 78:7314–7317.

Antoniades, H. N., C. D. Scher, and C. D. Stiles. 1979. Purification of platelet-derived growth factor. *Proc. Natl. Acad. Sci. USA.* 76:1809–1813.

Barrett, T. B., and E. P. Benditt. 1987. Sis (platelet-derived growth factor B chain) gene transcript levels are elevated in human atherosclerotic lesions compared to normal artery. *Proc. Natl. Acad. Sci. USA.* 84:1099–1103.

Barrett, T. B., and E. P. Benditt. 1988. Platelet-derived growth factor gene expression in human atherosclerotic plaques and normal artery wall. *Proc. Natl. Acad. Sci. USA.* 85:2810–2814.

Bedard, P.-A., D. Alcorn, D. L. Simmons, K.-C. Luk, and R. L. Erickson. 1987. Constitutive expression of a gene encoding a polypeptide homologous to biologically active human platelet protein in Rous sarcoma virus-transformed fibroblasts. *Proc. Natl. Acad. Sci. USA.* 84:6715–6719.

Betsholtz, C., B. Westermark, B. Ek, and C. H. Heldin. 1984. Co-expression of PDGF-like growth factor and PDGF receptors in human osteosarcoma cell line: implications for autocrine receptor activation. *Cell.* 39:447–457.

Boothroa, D. T., C. C. Morton, S. H. Orkin, and T. Collins. 1988. Platelet-derived growth factor A chain: Gene structure, chromosomal location, and basis for alternative mRNA splicing. *Proc. Natl. Acad. Sci. USA.* 85: 1492–1496.

Bywater, M., F. Rorsman, E. Bongcam-Rudloff, G. Mark, A. Hammacher, C-H. Heldin, B. Westermark, and C. Betsholtz. 1988. Expression of recombinant platelet-derived growth factor A and B chain homodimers in Rat-1 cells and human fibroblasts reveals differences in protein processing and autocrine effects. *Mol. Cell. Biol.* 8:2753–2762.

Cochran, B. H., A. C. Reddel, and C. D. Stiles. 1983. Molecular cloning of gene sequences regulated by platelet-derived growth factor. *Cell.* 33:939–947.

Collins, T., D. Ginsburg, J. M. Boss, S. H. Orkin, and J. S. Pober. 1985. Cultured human endothelial cells express platelet-derived growth factor B chain: cDNA cloning and structural analysis. *Nature (Lond.).* 316:748–750.

Collins, T., D. T. Boothroa, and S. H. Orkin. 1987. Alternative splicing affects function of encoded platelet-derived growth factor A chain. *Nature (Lond.).* 328:621–624.

DiCorleto, P. E. 1984. Cultured endothelial cells produce multiple growth factors for connective tissue cells. *Exp. Cell Res.* 153:167–172.

Dumont, J. N. 1972. Oogenesis in Xenopus laevis (Daudin). I. Stage of oocyte development in laboratory maintained animals. *J. Morphol.* 136:153–180.

Gimbrone, M. A., A. Hammacher, C. Betsholtz, B. Westermark, and C-H. Heldin. 1987. Cultured human endothelial cells express platelet-derived growth factor A chain. *Am. J. Physiol.* 126:7–12.

Gospodarowicz, D., K. D. Brown, C. R. Birdwell, and B. R. Zetter. 1978. Control of proliferation of human vascular endothelial cells: characterization of the response of human umbilical vein endothelial cells to fibroblast growth factor, epidermal growth factor, and thrombin. *J. Cell Biol.* 77:774–788.

Grotendorst, G. R. 1984. Alteration of the chemotactic response of NIH/3T3 cells to PDGF growth factors, transformation and tumor promoters. *Cell.* 36:279–285.

Grotendorst, G. R. 1987. Spectrophotometric assay for the quantitation of cell migration in the Boyden chamber chemotaxis assay. *Methods Enzymol.* 147:144–152.

Grotendorst, G. R., and G. R. Martin. 1986. Cell movements in wound-healing and fibrosis. *Rheumatology.* 10:385–403.

Grotendorst, G. R., H. E. Seppa, H. K. Kleinman, and G. R. Martin 1981. Attachment of smooth muscle cells to collagen and their migration toward platelet derived growth factor. *Proc. Natl. Acad. Sci. USA.* 78:3669–3672.

Heldin, C.-H., B. Westermark, and A. Wasteson. 1979 Purification of platelet-derived growth factor and partial characterization. *Proc. Natl. Acad. Sci. USA.* 76:3722–3726.

Heldin, C.-H., B. Westermark, and A. Wasteson. 1981. Specific receptors for platelet-derived growth factor on cells derived from connective tissue and glia. *Proc. Natl. Acad. Sci. USA.* 78:3364–3368.

Heldin, C.-H., A. Johnsson, S. Wennergren, C. Wennergren, C. Wernstedt, C. Betsholtz, and B. Westermark. 1986. A human osteosarcoma cell line secretes a growth factor structurally related to a homodimer of PDGF A chains. *Nature (Lond).* 319:511–514.

Jaffe, E. A. 1984. Biology of Endothelial Cells. Martinus-Nijhoff, Boston.

Jaffe, E. A. 1987. Cell biology of endothelial cells. *Human Pathol.* 18:234–239.

Jaffe, E. A., R. L. Nachman, C. G. Becker, and R. Minick. 1973. Culture of human endothelial cells derived from umbilical veins. *J. Clin. Invest.* 52:2745–2756.

Kack, P. J., S. D. Houser, G. Krivi, K. Sanzo, T. Warren, J. Edeer, and D. T. Connolly. 1989. Vascular Permeability Factor, an endothelial cell mitogen related to PDGF. *Science (Wash. DC).* 246:1309–1312.

Kohler, N., and A. Lipton. 1974. Platelets as a source of fibroblast growth-promoting activity. *Exp. Cell. Res.* 87:297–301.

Laemmli, U. K. 1970. Cleavage of structural proteins during assembly of the head of bacteriophage T4. *Nature (Lond.).* 227:680–685.

Leung, D. W., G. Cachianes, W-J. Kuang, D. V. Goeddel, and N. Ferrara. 1989. Vascular Endothelial Growth Factor is a secreted angiogenic mitogen. *Science (Wash. DC).* 246:1306–1309.

Majesty, M. W., E. P. Benditt, and S. M. Schwartz. 1988. Expression and developmental control of platelet-derived growth factor A-chain and B-chain/sis genes in rat cortic smooth muscle cells. *Proc. Natl. Acad. Sci. USA.* 85: 1524–1528.

Martinet, Y., P. B. Bitterman, J. F. Marnex, G. R. Grotendorst, G. R. Martin, and R. G. Crystal. 1986. Activated human monocytes express the c-sis proto-oncogene and release a mediator showing PDGF-like activity. *Nature (Lond.).* 319:158–160.

Oquendo, P., J. Alberta, D. Wen, J. L. Greycar, R. Deryack, and C. D. Stiles. 1989. The platelet-derived growth factor-inducible KC gene encodes a secretory protein related to platelet α-granule proteins. *J. Biol. Chem.* 264: 4133–4137.

Pencev, D., and G. R. Grotendorst. 1988. Human peripheral blood monocytes secrete a unique form of PDGF. *Oncogene Res.* 3:333–342.

Robbins, K. C., H. N. Antoniades, S. G. Devare, M. W. Hunkapiller, and S. A. Aaronson. 1983. Structural and immunological similarities between simian sarcoma virus gene product(s) and human platelet-derived growth factor. *Nature (Lond.).* 305:605–608.

Rollins, B. J., and C. D. Stiles. 1989. Serum-inducible genes. *Adv. Cancer Res.* 53:1–32.

Rollins, B. J., E. D. Morrison, and C. D. Stiles. 1988. Cloning and expression of JE, a gene inducible by platelet-derived growth factor and whose product has cytokine-like properties. *Proc. Natl. Acad. Sci. USA.* 85:3738–3742.

Ross, R., J. Glomset, B. Kariya, and L. Harker. 1974. A platelet dependent serum factor that stimulates the proliferation of arterial smooth muscle cells in vitro. *Proc. Natl. Acad. Sci. USA.* 71:1207–1210.

Ross, R., E. W. Raines, and D. F. Bowen-Pope. 1986. The biology of platelet-derived growth factor. *Cell.* 46:155–169.

Sambrook, J., E. F. Fritsch, and T. Maniatis. 1989. Molecular Cloning: A Laboratory Manual. Second edition. Cold Spring Harbor Laboratory, Cold Spring Harbor, NY. 545 pp.

Shimokado, K., E. W. Raines, D. K. Madtes, T. B. Barrett, E. P. Benditt, and R. Ross. 1985. A significant part of macrophage-derived growth factor consists of at least two forms of PDGF. *Cell.* 43:277–286.

Simmons, D. L., D. B. Levy, Y. Yannoni, and R. L. Erikson. 1989. Identification of a phorbol ester-repressible v-src-inducible gene. *Proc. Natl. Acad. Sci. USA.* 86:1178–1182.

Sirtras, N. M., E. Soriban, P. Pantazis, B. Zetter, and H. N. Antoniades. 1987. Human iliac artery endothelial cells express both gene encoding the chains of platelet-derived growth factor (PDGF) and synthesize PDGF-like mitogen. *J. Cell. Physiol.* 132:376–380.

Stroobant, P., and M. D. Waterfield. 1984. Purification and properties of porcine platelet-derived growth factor. *EMBO (Eur. Mol. Biol. Organ.) J.* 3:2963–2967.

Studier, F. W., A. H. Rosenburg, J. J. Dunn, and J. W. Dubendorff. 1990. Use of T7 RNA polymerase to direct expression of cloned genes. *Methods Enzymol.* 185:60–89.

Tischer, E., D. Gospodoarwicz, R. Mitchell, M. Silva, J. Schilling, K. Lau, T. Crisp, J. C. Fiddes, and J. A. Abraham. 1989. Vascular endothelial growth factor: a new member of the platelet-derived growth factor gene family. *Biochem. Biophys. Res. Commun.* 165:1198–1206.

Tong, B. D., D. E. Auer, M. Joye, J. M. Daplow, G. Rieca, E. McConathy, W. Drohan, and T. F. Neuel. 1987. cDNA clones reveal differences between human glial and endothelial cell platelet-derived growth factor A-chains. *Nature (Lond.).* 328:619–621.

Towbin, H., T. Smehelin, and J. Gorden. 1979. Electrophoretic transfer of protein from polyacrylamide gels to nitrocellulose sheets: procedure and some application. *Proc. Natl. Acad. Sci. USA.* 76:4353–4356.

Wallace, R. A., D. W. Jared, and J. N. Dumont. 1973. Protein incorporation by isolated amphibian oocytes. *Exp. Zool.* 184:321–334.

Cell Biology

Contents:
The Journal of Cell Biology

Volume 114, Number 6, September 1991

Contents continued on reverse of this cover

Cover picture: Multiple exposure, moving film, photograph of a demembranated sea urchin spermatozoon reactivated with ATP. Gold beads attached to the exposed outer doublet microtubules of the sperm flagellum allow direct measurement of sliding between the microtubules as the flagellum bends. See related article in this issue by Brokaw, 1201–1215.

Docket No.: PF-0501-1 DIV
USSN: 09/866,356
Attachment B

1: Mol Cell Biol. 1989 Dec;9(12):5718-21. Related Articles, Links

Definition of an Ets1 protein domain required for nuclear localization in cells and DNA-binding activity in vitro.

Boulukos KE, Pognonec P, Rabault B, Begue A, Ghysdael J.

Institut National de la Sante de la Recherche Medicale U186/Centre National de la Recherche Scientifique UA 041160, Institut Pasteur, Lille, France.

Ets1 and Ets2 are nuclear phosphoproteins which bind to DNA in vitro and share two domains of strong identity. Deletion analyses of each of these conserved regions in Ets1 demonstrated that integrity of the carboxy-terminal domain, also conserved in the more distantly related elk and erg gene products, is essential for both nuclear targeting and DNA-binding activity in vitro.

PMID: 2555704 [PubMed - indexed for MEDLINE]

Display Abstract Show: 20 Sort Send to Text

Entrez PubMed
Overview
Help | FAQ
Tutorial
New/Noteworthy
E-Utilities

PubMed Services
Journals Database
MeSH Database
Single Citation Matcher
Batch Citation Matcher
Clinical Queries
LinkOut
Cubby

Related Resources
Order Documents
NLM Gateway
TOXNET
Consumer Health
Clinical Alerts
ClinicalTrials.gov
PubMed Central

Privacy Policy

Write to the Help Desk
NCBI | NLM | NIH
Department of Health & Human Services
Freedom of Information Act | Disclaimer

Oct 2 2003 18:06:20

THE JOURNAL OF BIOLOGICAL CHEMISTRY
Vol. 275, No. 2, Issue of January 14, pp. 1216–1225, 2000
Printed in U.S.A.

PDEF, a Novel Prostate Epithelium-specific Ets Transcription Factor, Interacts with the Androgen Receptor and Activates Prostate-specific Antigen Gene Expression*

(Received for publication, September 2, 1999, and in revised form, October 28, 1999)

Peter Oettgen‡, Eduardo Finger‡, Zijie Sun§, Yasmin Akbarali‡, Usanee Thamrongsak‡, Jay Boltax‡, Franck Grall‡, Antoinise Dube‡, Avi Weiss‡, Lawrence Brown¶, Gary Quinn‖, Koen Kas‡ **, Greg Endress‡‡, Charles Kunsch‡‡, and Towia A. Libermann‡ §§

From the ‡New England Baptist Bone and Joint Institute, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts 02115, the §Department of Surgery and Genetics, Stanford University School of Medicine, Stanford, California 94305, the ¶Division of Pathology, Beth Israel Deaconess Medical Center and Harvard Medical School, Boston, Massachusetts 02215, the ‖Department of Biology, University of York, P.O. Box 373, York YO1 5YW, United Kingdom, and ‡‡Human Genome Sciences, Inc., Rockville, Maryland 20850

Prostate cancer, the most frequent solid cancer in older men, is a leading cause of cancer deaths. Although proliferation and differentiation of normal prostate epithelia and the initial growth of prostate cancer cells are androgen-dependent, prostate cancers ultimately become androgen-independent and refractory to hormone therapy. The prostate-specific antigen (PSA) gene has been widely used as a diagnostic indicator for androgen-dependent and -independent prostate cancer. Androgen-induced and prostate epithelium-specific PSA expression is regulated by a proximal promoter and an upstream enhancer via several androgen receptor binding sites. However, little progress has been made in identifying androgen-independent regulatory elements involved in PSA gene regulation. We report the isolation of a novel, prostate epithelium-specific Ets transcription factor, PDEF (prostate-derived Ets factor), that among the Ets family uniquely prefers binding to a GGAT rather than a GGAA core. PDEF acts as an androgen-independent transcriptional activator of the PSA promoter. PDEF also directly interacts with the DNA binding domain of androgen receptor and enhances androgen-mediated activation of the PSA promoter. Our results, as well as the critical roles of other Ets factors in cellular differentiation and tumorigenesis, strongly suggest that PDEF is an important regulator of prostate gland and/or prostate cancer development.

Prostate cancer, the most common solid cancer in older men, is one of the most frequent causes of cancer deaths. The lack of effective therapies for advanced prostate cancer reflects in part the lack of knowledge about the molecular mechanisms involved in the development and progression of this disease (1, 2). In particular, little is known about the mechanisms that trigger the conversion of an initially androgen-dependent cancer to androgen independence (1, 2). Prostate cancer development can be divided into different steps of epithelial cell transformation (3). Proliferation within the normal and hyperplastic prostate epithelium is restricted to the basal cells and is regulated by autocrine or paracrine growth factors. Differentiation from basal cells to secretory luminal epithelial cells is androgen-dependent, and androgen remains crucial for the initial growth of prostate cancer cells (3, 4). Nevertheless, prostate cancers ultimately become androgen-independent and refractory to hormone therapy. PSA[1] has been used widely as a diagnostic indicator for prostate cancer (5). PSA is expressed by normal and cancerous luminal epithelial cells of the prostate, and its expression is under the control of androgens acting through the AR (5). However, even in hormone-refractory prostate cancer, PSA is expressed, suggesting an androgen-independent component in PSA regulation as well (6). PSA gene expression is regulated by the proximal promoter and a strong upstream enhancer region (7–9). Both regulatory regions contain binding sites for AR and are essential for androgen-induced transcriptional activation of the PSA gene. However, little progress has been made in identifying androgen-independent regulatory elements involved in PSA gene regulation.

We report here the characterization of a novel prostate epithelial-specific Ets transcription factor, PDEF, that is involved in PSA gene regulation and acts as a co-regulator of AR. Ets factors play a crucial role in the regulation of genes involved in hematopoiesis, angiogenesis, organogenesis, and specification of neuronal connectivity (10–12), and several distinct chromosomal translocations involving various Ets factors have been discovered in human cancer (10–12). The recent isolation of three epithelial-specific Ets factors, ESE-1 (ESX/ELF3/ERT/JEN), ESE-2 (ELF5), and ESE-3 (EHF), has demonstrated the relevance of Ets factors in epithelial cells (13–21). Our results now support the notion that PDEF is involved in prostate epithelium-specific gene expression and possibly in prostate cancer development or progression.

MATERIALS AND METHODS

Cell Culture—Human foreskin keratinocytes, HaCAT (keratinocyte line), HEK293 (fetal epithelial kidney), C-33A (cervical carcinoma), HeLa (cervical carcinoma), H157 (large cell lung carcinoma), H249 (small cell lung carcinoma), HUVEC (endothelial), U-937 (monocytes),

* This work was supported by a CaP CURE award (to T. A. L.) and by National Institutes of Health Grant KO8/CA 71429 (to P. O.) The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

The nucleotide sequence(s) reported in this paper has been submitted to the GenBank™/EBI Data Bank with accession number(s) AF071538.

** Postdoctoral student of the Fonds Wettenschappelidk Onderzoek.

§§ To whom correspondence should be addressed: New England Baptist Bone and Joint Inst., Dept. of Medicine, Beth Israel Deaconess Medical Center, 4 Blackfan Circle, Boston, MA 02115. Tel.: 617-667-3393; Fax: 617-975-5299; E-mail: tliberma@caregroup.harvard.edu.

[1] The abbreviations used are: PSA, prostate-specific antigen; AR, androgen receptor; DHT, dihydrotestosterone; PDEF, prostate-derived Ets factor; EMSA, electrophoretic mobility shift assay; EST, expressed sequence tag; RT, reverse transcriptase; PCR, polymerase chain reaction; GST, glutathione *S*-transferase; kb, kilobase pair(s); MAP, mitogen-activated protein; GAPDH, glyceraldehyde-3-phosphate dehydrogenase.

human synovial fibroblasts, and human chondrocytes were grown as described (15, 22). LNCaP (prostate cancer) cells were grown in T-medium (Life Technologies, Inc.) with 10% fetal calf serum.

Isolation of PDEF cDNA—A human expressed sequence tag (EST) from a subtracted prostate benign hyperplasia cDNA library with significant homology to the ETS domain was identified in the Human Genome Sciences EST data base using the tBLASTN program (NCBI). To isolate the entire coding region for PDEF, we performed the rapid amplification of cDNA ends method using human adult prostate Marathon Ready cDNA (CLONTECH) as described (15, 23).

RT-PCR Analysis and Northern Blot Analysis—mRNAs were isolated as described (24, 25). Northern blots containing poly(A)$^+$ mRNA (CLONTECH) were hybridized with random prime-labeled PDEF, ESE-1, and GAPDH cDNA probes in QuickHyb solution (Stratagene) (15).

cDNAs were generated from 1 μg of mRNA. RT-PCR reactions were performed and analyzed as described (15).

The sequences of the PDEF primers were: sense, 5′-GACCAGTGAGGAGAGCTGGACCGA-3′; antisense, 5′-TGACCTTGGGCTCTGGAAGGTCAG-3′ with an expected size of 473 base pairs.

The sequences of the primers for GAPDH were: sense, 5′-CAAAGTTGTCATGGATGACC-3′; antisense, 5′-CCATGGAGAAGGCTGGGG-3′ with an expected amplification product of 200 base pairs.

Expression Vector and Luciferase Reporter Gene Constructs—The PSA promoter containing 7 kb of 5′ upstream region was inserted into the pGL3 luciferase vector (pGL3/PSA). The PSP954 promoter and the CRISP-3 promoter were inserted into the pGL3 promoter as described (21). The full-length PDEF cDNA as well as other Ets factor cDNAs as described under "Results" were inserted into the *Eco*RI site of the pCI (Promega) eukaryotic expression vector downstream of the cytomegalovirus and T7 promoters (pCI/PDEF). Coupled *in vitro* transcription/*in vitro* translation reactions of full-length PDEF and other Ets factors were performed (Promega) as described (26). The human AR expression vector pAR0 was kindly provided by Dr. Albert Brinkmann (Erasmus University, Rotterdam, The Netherlands) (27).

Electrophoretic Mobility Shift Assay (EMSA)—EMSA were performed as described (22, 26) using 2 μl of *in vitro* translation product and 0.2 ng of ^{32}P-labeled double-stranded oligonucleotide probes (20000 cpm) and run on 4% polyacrylamide gels, containing as buffer 0.5× TGE as described (15).

Oligonucleotides used as probes are as follows.

1) PSA promoter oligonucleotide WT Ets site A (−268):
5′-TCGAGGCCCAGGATGAAACAG-3′
3′-CCGGGTCCTACTTTGTCAGCT-5′
2) PSA promoter oligonucleotide WT Ets site B (−5494):
5′-TCGAGAGCCAGGATGGTCTCG-3′
3′-CTCGGTCCTACCAGAGCAGCT-5′
3) PSA promoter oligonucleotide WT Ets site C (−5215):
5′-TCGAGTGGCAGGATATTCCAG-3′
3′-CACCGTCCTATAAGGTCAGCT-5′
4) PSA promoter oligonucleotide WT Ets site D (−4039):
5′-TCGAGTGCAAGGATGCCTGCG-3′
3′-CACGTTCCTACGGACGCAGCT-5′
5) PSA promoter oligonucleotide WT Ets site E (−3848):
5′-TCGAGAAGCAGGATGTGATAG-3′
3′-CTTCGTCCTACACTATCAGCT-5′
6) PSA promoter oligonucleotide WT Ets site F (−3536):
5′-TCGAGCCTCAGGAAGAAGGTG-3′
3′-CGGAGTCCTTCTTCCACAGCT-5′
7) PSA promoter oligonucleotide WT Ets site G (−2174):
5′-TCGAGACACCGGATCCATGAG-3′
3′-CTGTGGCCTAGGTACTCAGCT-5′
8) PSA promoter oligonucleotide WT Ets site H (−2129):
5′-TCGAGTTCCAGGATGAATCCG-3′
3′-CAAGGTCCTACTTAGGCAGCT-5′
9) PSA promoter oligonucleotide WT Ets site I (−4156):
5′-TCGACACTCTGGAGGAACATATTGTATCGAC-3′
3′-GTGAGACCTCCTTGTATAACATAGCTGAGCT-5′
10) PSA promoter oligonucleotide WT Ets site J (−149):
5′-TCGAGCAGAGCTGTGGAAGGGGAGGG-3′
3′-CGTCTCGACACCTTCCCCTCCCAGCT-5′
11) PSA promoter oligonucleotide WT Ets site K (−45):
5′-TCGACGTCCTGGGGAATGAAGGTTTC-3′
3′-GCAGGACCCCTTACTTCCAAAGAGCT-5′
12) PSA promoter oligonucleotide mutant Ets site E (GGAA):
5′-TCGAGAAGCAGGAAGTGATAG-3′
3′-CTTCGTCCTTCACTATCAGCT-5′

DNA Transfection Assays—Co-transfections of 3 × 10^5 CV-1 (green monkey kidney) cells were carried out with 3.5 μg of pGL3/PSA DNA, 2.5 μg of pAR0 AR expression vector, and 1.5 μg of pCI/PDEF DNA using 12.5 μl of LipofectAMINE (Life Technologies, Inc.) as described (15). Cells grown in medium containing charcoal-stripped serum for 24 h were incubated in the absence or presence of 10^{-8} M dihydrotestosterone (DHT) for an additional 16 h and harvested for luciferase activity as described (28). Co-transfections of 3 × 10^5 LNCaP cells were carried out with 1.75 μg of pGL3/PSA DNA and 0.75 μg of pCI/PDEF DNA using 4 μl LipofectAMINE Plus (Life Technologies, Inc.) as described (15). Cells were grown in medium containing charcoal-stripped serum and assayed 36 h later for luciferase activity as described (28). Transfections were performed independently in triplicate and repeated three times with two different plasmid preparations with similar results. Cotransfection of a second plasmid for determination of transfection efficiency was omitted because potential artifacts with this technique have been reported (29).

In Situ Hybridization—Tissues fixed in 4% paraformaldehyde in phosphate-buffered saline, pH 7.4, for 2–4 h at 4 °C were transferred to 30% sucrose in phosphate-buffered saline overnight at 4 °C, frozen in OCT compound (Miles Diagnostics, Elkhart, IN), and stored at −70 °C. *In situ* hybridization was performed on 6-μm frozen sections with ^{35}S-labeled riboprobes as described (15).

GST Pull-down Assay—A series of GST-AR fusion proteins were generated by PCR with specific primers to contain in frame restriction enzyme sites (30) and sequenced to confirm that there were no mutations introduced by the PCR. GST fusion proteins were prepared as described before (31). 5 μl of [^{35}S]methionine labeled *in vitro* translated full-length PDEF protein was incubated with equal amounts of GST-AR fusion proteins or GST on agarose beads in 200 μl of NETN (0.5% Nonidet P-40, 1 mM EDTA, 20 mM Tris-HCl, pH 8.0, 100 mM NaCl) for 3 h at 4 °C with gentle shaking. PDEF protein that bound to the GST-AR fusion proteins was eluted after three washings with NETN buffer and analyzed on a 12% SDS-polyacrylamide gel.

RESULTS

Isolation of PDEF and Sequence Comparison with Other Members of the Ets Family—We have isolated full-length cDNA clones from human prostate encoding a novel member of the Ets transcription factor/oncogene family, PDEF (Fig. 1*A*). Sequence analysis revealed an open reading frame encoding a 335-amino acid protein with a predicted molecular mass of 37.5 kDa. The deduced amino acid sequence of PDEF predicts a protein rich in glutamic acid (8%), alanine (8%), serine (11%), leucine (10%), and proline (8%). Two putative PEST domains, common to rapidly degraded proteins, are located between amino acids 19 and 53 and between 161 and 186. Potential phosphorylation sites present in PDEF include a protein kinase C site, two AKT phosphorylation sites, two tyrosine kinase phosphorylation sites, and eight c-Jun NH_2-terminal kinase/p38/extracellular signal-regulated kinase kinase phosphorylation sites. Five of these MAP kinase phosphorylation sites are clustered at the amino terminus of PDEF, and one of these sites conforms to the optimal MAP kinase phosphorylation site P*X*(S/T)P (Fig. 1*A*). One of the tyrosine kinase phosphorylation sites contains the consensus sequence for interaction with phosphoinositol 3-kinase, Y*XX*M (32).

Ets transcription factors share a highly conserved DNA binding domain, the ETS domain (10–12). Alignment of the carboxyl-terminal ETS domain of PDEF with that of other Ets family members reveals highest homology to one particular subclass of Ets factors that includes *Drosophila* D-ets-4 (75%) and sea urchin SpETS4 (76%) (Fig. 1*B*). SpEts4 plays a role in establishing the animal-vegetal axis of the sea urchin embryo at late cleavage/early blastula stages (33). D-Ets-4 is expressed highest in the pole cells, suggesting an involvement in germ line differentiation (34). Upstream of the Ets domain, PDEF contains a region with significant homology to the Pointed domain present in several other members of the Ets family including ESE-1, ESE-2, ESE-3, Tel, Tel-2, yan, pointed, Ets-1, Ets-2, fli-1, erg-2, Elg, and GABP-α and weakly homologous to the SAM protein-protein interaction domain of polycomb proteins and Eph receptors (35) (Fig. 1*C*). In contrast to all other Ets factors where the Pointed domain is directly at the amino

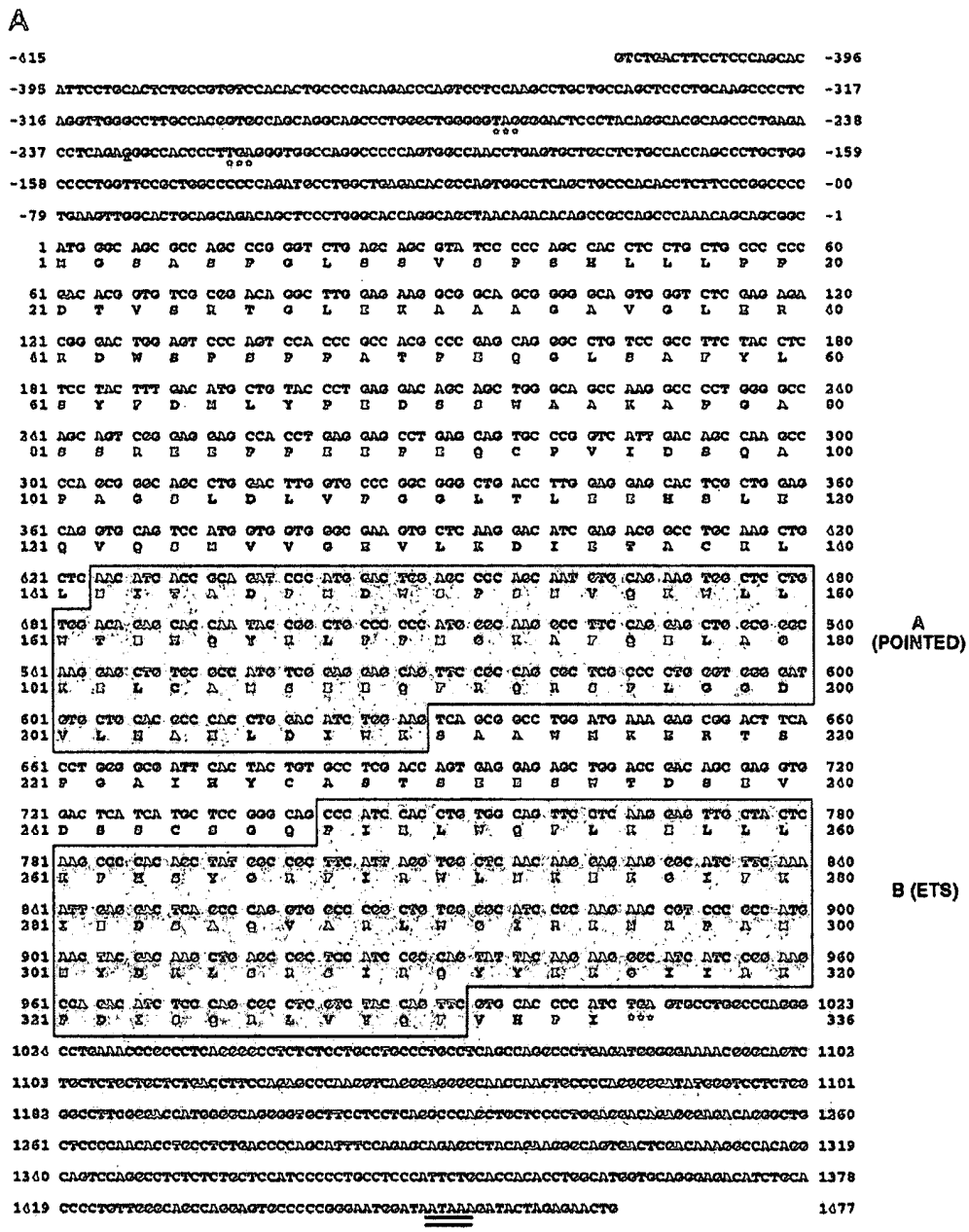


FIG. 1. **Sequence of PDEF.** *A*, nucleotide and amino acid sequence of PDEF. The deduced amino acid sequence of PDEF is indicated *below* the first nucleotide of each codon, and the termination codon is marked with *three asterisks*. The Pointed domain and the ETS domain are *shaded*, *boxed*, and marked on the *right*. Both upstream termination codons in frame with the reading frame are also indicated by *asterisks*. The polyadenylation signal is *underlined*. *B*, comparison of the Ets domain of PDEF with all known members of the Ets gene family. Percentage of identity of the ETS domain with PDEF is indicated on the *right side*. *Shaded* amino acids denote amino acid identity with PDEF. Gaps are introduced to optimize alignment. The proteins examined are indicated on the *left side*. Tentative α helices and β sheets as predicted by NMR and x-ray crystallography for some of the Ets domains are indicated above the sequence by *boxes* or *arrows*, respectively. *C*, comparison of the Pointed domain of PDEF with all known members of the ETS gene family. Shown are the homologies between the PDEF Pointed domain and the corresponding regions present in a subset of the ETS family. Only human and *Drosophila* ETS factors are included for simplicity. The *consensus line* indicates which residues are highly conserved.

terminus and the Ets domain at the opposite carboxyl terminus, PDEF has the Pointed domain immediately upstream of the Ets domain.

PDEF Expression Is Restricted to the Epithelial Layers of the Prostate and to a Lower Extent to Other Hormone-regulated Tissues—Expression of PDEF was analyzed by Northern blot hybridization of poly(A)$^+$ mRNAs derived from human adult and fetal tissues, indicating the presence of one predominant PDEF transcript of ~2.0 kb (Fig. 2*A*). PDEF is almost exclusively expressed in the prostate gland. Low levels of PDEF transcripts were also found in ovary (Fig. 2*A*) and in mammary and salivary gland (data not shown), all tissues regulated by androgens or steroids. No PDEF transcripts were detected in brain, heart, gastrointestinal tract, liver, spleen, thymus, kidney, lung, and skeletal muscle or in fetal tissues. These results suggest that PDEF is expressed in a very restricted set of primarily hormone-regulated epithelial tissues with particularly high expression in prostate, an expression pattern strikingly different from other Ets factors such as ESE-1 (Fig. 2*A*).

More detailed analysis of PDEF expression in different cell types by RT-PCR revealed that only the prostate epithelial cancer cell line, LNCaP, expressed PDEF mRNA, whereas

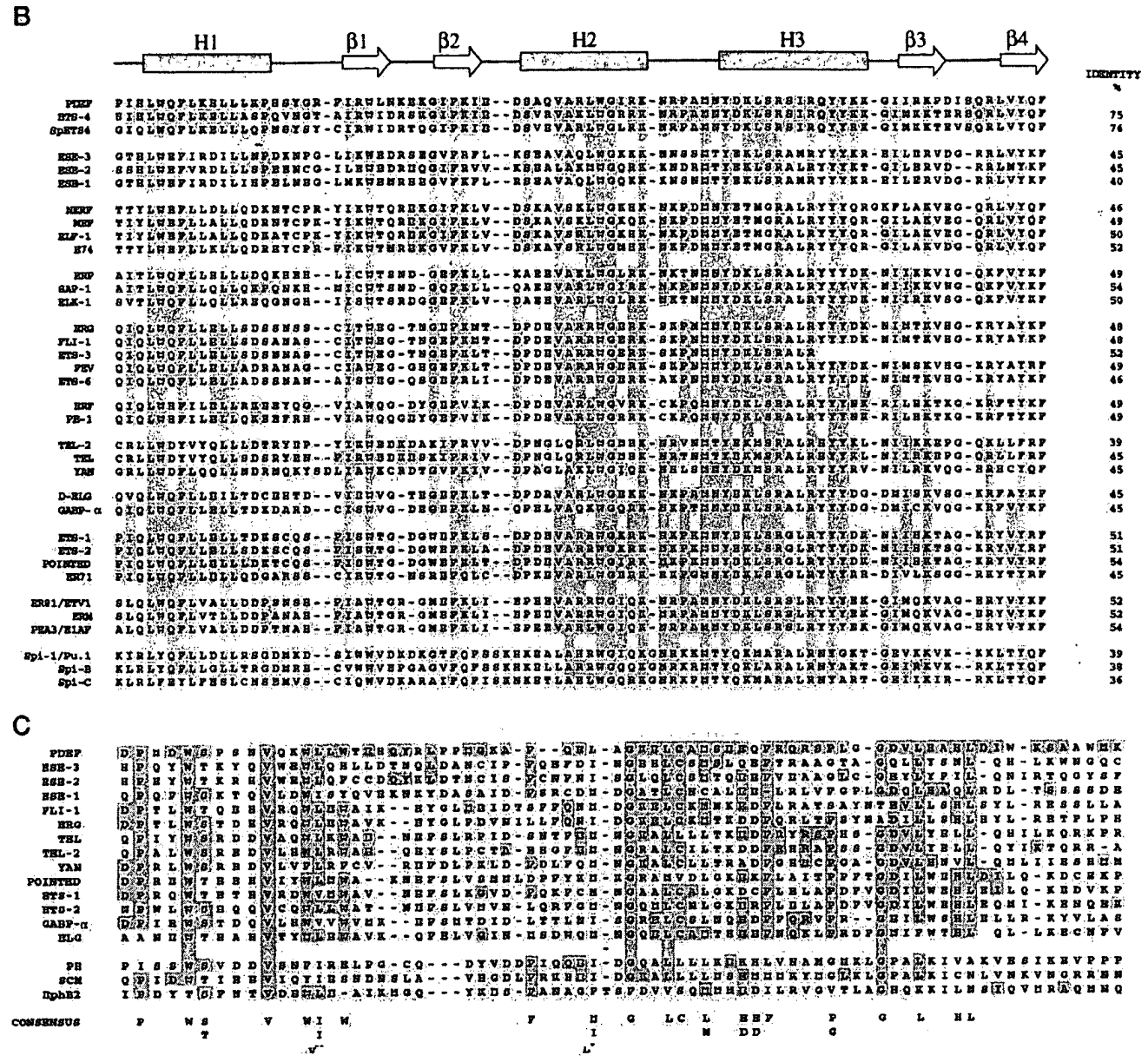


FIG. 1—*continued*

other epithelial cells and non-epithelial cells were completely devoid of PDEF mRNA (Fig. 2*B*). To further test the hypothesis that PDEF expression is restricted to epithelial cells within the prostate gland, we performed *in situ* hybridization on frozen sections of human prostate. Diffuse strong expression was only noted in luminal epithelium of the prostate, but not in other cell types (Fig. 2*C*). These results most vividly demonstrate that PDEF is exclusively expressed in epithelial cells and particularly strong in the prostate.

PDEF Binds Specifically to Several Sequences in the PSA Promoter/Enhancer—In the prostate, PDEF expression is restricted to luminal epithelial cells, the exact cells that express PSA. The PSA gene is regulated by an androgen-responsive promoter and an androgen-responsive upstream enhancer (Fig. 3*A*). At least eight high and low affinity androgen receptor binding sites have been identified in the promoter and enhancer region and implicated in androgen-mediated regulation of the PSA gene. Inspection of the PSA promoter and enhancer sequence revealed the existence of at least 11 putative Ets binding sites. Some of these sites are in close proximity to androgen receptor binding sites (Fig. 3*A*). To determine whether the PSA promoter contains binding sites for PDEF, we tested the ability of *in vitro* translated PDEF to bind specifically to oligonucleotides containing a variety of Ets related binding sites derived from the PSA promoter. EMSA analysis revealed that the PSA E site and G site oligonucleotides formed two major complexes with proteins present in the PDEF extract (Fig. 3*B*), which were not formed by the control extract (Fig. 3*B*). A third strong complex was nonspecific and was present in both PDEF and control reticulocyte lysates. Two additional PSA sites, site H and to a lesser extent site C (Fig. 3*C*), formed weak complexes with PDEF, whereas other potential Ets sites in the PSA promoter did not significantly interact with PDEF. Strikingly, the high affinity binding sites for PDEF diverge from the binding sites found for other Ets factors showing an apparent preference for GGAT *versus* GGAA in the core of the binding site (Fig. 3*C*). EMSA analysis of additional canonical Ets sites from other promoters did not reveal any significant PDEF binding,

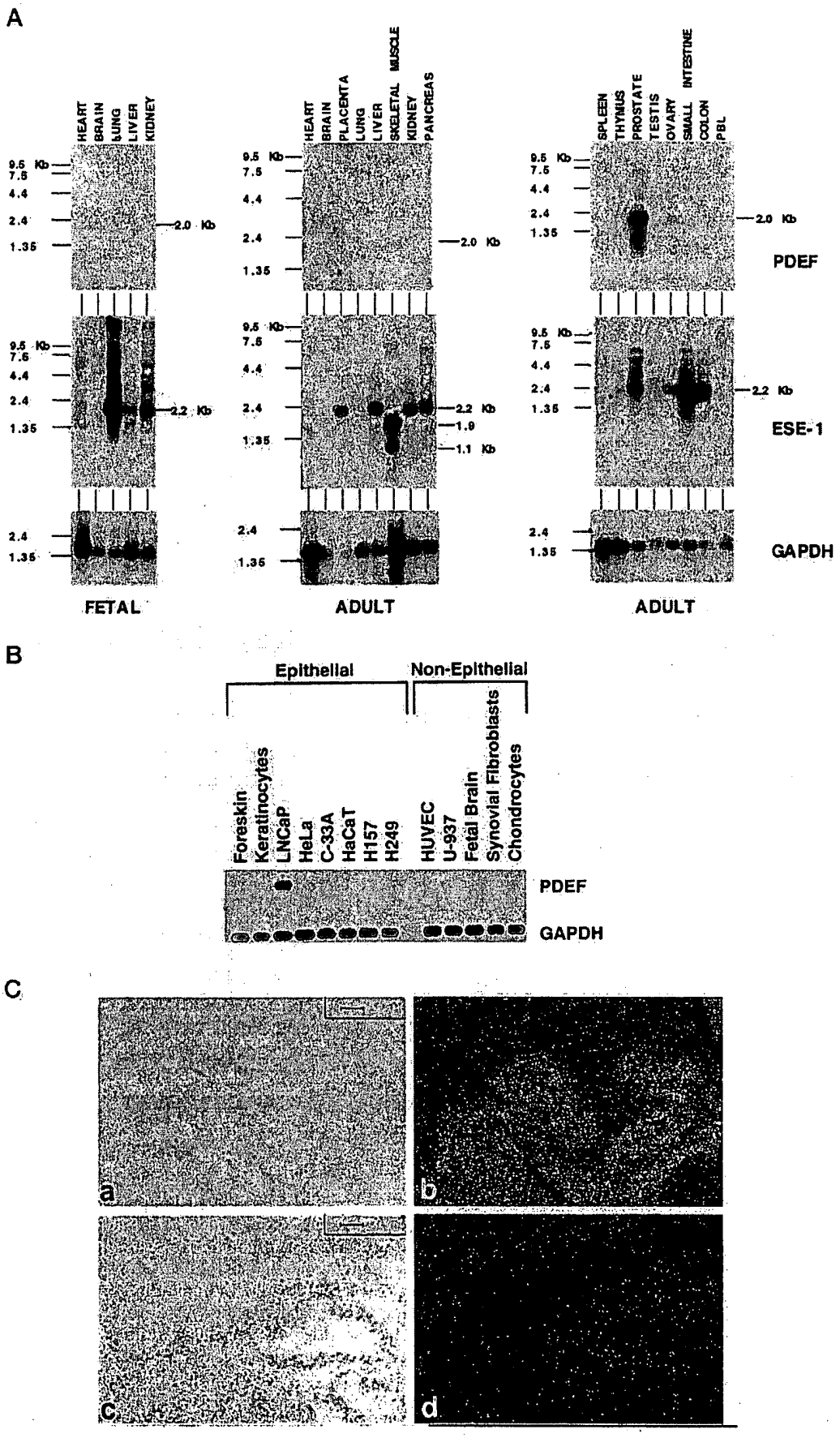
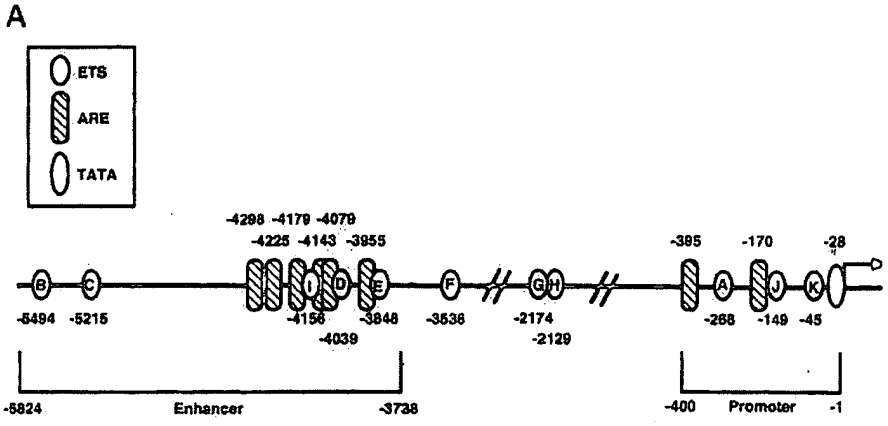
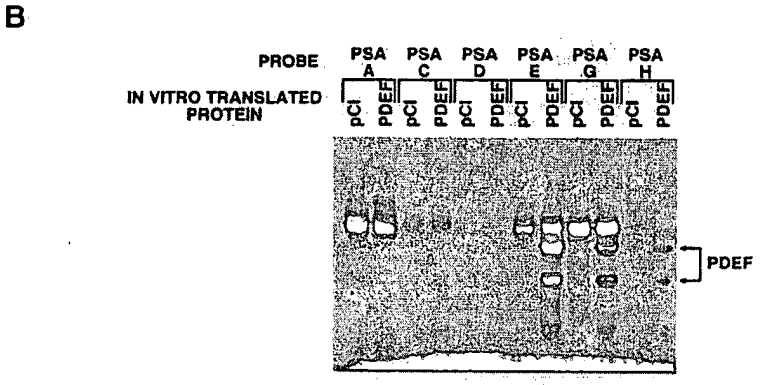


FIG. 2. **Expression of PDEF in different human fetal and adult tissues.** *A*, Northern blot analysis of poly(A)$^+$ mRNA from human fetal and adult tissues. The blots were sequentially probed with a PDEF (*upper panel*), ESE-1 (*middle panel*), and a GAPDH cDNA probe (*lower panel*) under





C

PSA PROMOTER	BINDING SITE	PDEF
PSA E	AGCAGGATGTGA	++++
PSA G	CACCGGATCCAT	++++
PSA H	TCCAGGATGAAT	++
PSA C	GGCAGGATATTC	+
PSA A	CCCAGGATGAAA	+/-
PSA D	GCAAGGATGCCT	+/-
PSA B	GCCAGGATGGTC	-
PSA F	CTCAGGAAGAAG	-
PSA I	TGGAGGAACATA	-
PSA J	CTGTGGAAGGGG	-
PSA K	CTGGGGAATGAA	-
CONSENSUS	NGCAGGATGTAT	
	A C CCGA	
	C AATC	

FIG. 3. **PDEF binds to sequences in the PSA promoter.** *A*, schematic diagram of the human PSA promoter/enhancer region. The locations of putative regulatory elements including the androgen receptor binding sites and the Ets binding sites are indicated. The *letters* within the *symbols* correspond to the Ets sites as described under "Material and Methods." *B*, several PSA promoter Ets sites are targets for PDEF. Panel shows DNA binding of PDEF to PSA promoter Ets sites (see "Material and Methods" and *panel C*), and DNA binding of full-length PDEF in an EMSA using synthetic oligonucleotides coding for the PSA promoter Ets sites using either reticulocyte lysate programmed with the empty pCI expression vector (*pCI*) or reticulocyte lysate programmed with the PDEF pCI expression vector (*PDEF*). *C*, relative DNA binding affinity of PDEF toward a variety of putative Ets binding sites within the PSA promoter as measured by EMSA analysis. A putative high affinity consensus PDEF binding site is shown at the *bottom*.

suggesting strong DNA binding selectivity for PDEF.[2]

The DNA Binding Specificity of PDEF Is Uniquely Distinct from Other Members of the Ets Family—To directly assess the relative DNA binding specificity and affinity of PDEF and other members of the Ets family, we compared the ability of PDEF and several other Ets factors to bind to the GGAT containing PSA promoter E site and to the same site containing a single nucleotide change to GGAA in the core of the binding site. PDEF bound with high affinity only to the GGAT-containing oligonucleotide, but not to the GGAA-containing oligonucleotide (Fig. 4). Some low affinity binding of ESE-3 to the GGAT oligonucleotide was also observed. In striking contrast

[2] P. Oettgen, E. Finger, Z. Sun, Y. Akbarali, U. Thamrongsak, J. Boltax, F. Grall, A. Dube, A. Weiss, L. Brown, G. Quinn, K. Kas, G. Endress, C. Kunsch, and T. A. Libermann, manuscript in preparation.

stringent conditions. *B*, expression of PDEF in various cell types (see "Materials and Methods") by RT-PCR analysis of poly(A)$^+$ mRNA from indicated cells using primers specific for PDEF (*upper panel*) or GAPDH (*lower panel*). *C*, *in situ* hybridization studies. Panel shows paired bright field (*a* and *c*) and corresponding polarized fluorescence (*b* and *d*) photomicrographs. Intense labeling of prostate epithelium in normal prostate is seen with antisense probe to PDEF mRNA (*a* and *b*). No labeling is seen with control sense probe (*c* and *d*).

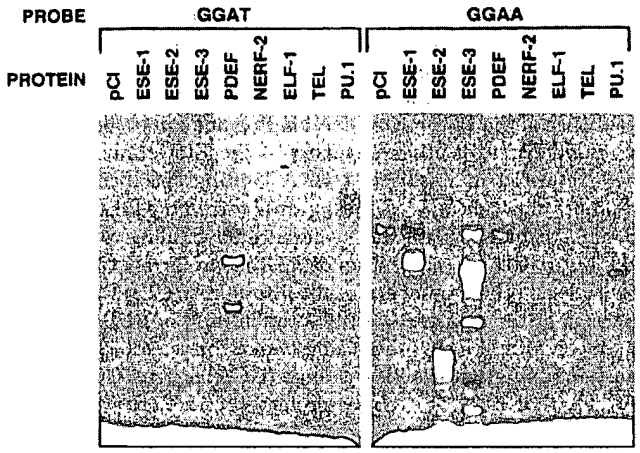


FIG. 4. **PDEF, in contrast to other Ets factors, preferentially binds to GGAT containing DNA sequences.** Figure shows binding of full-length *in vitro* translated ETS factors as indicated above the figure in an EMSA to synthetic oligonucleotides coding for either the wild type PSA promoter E Ets site (GGAT) containing GGAT in the core of the binding site or to a mutant PSA promoter E Ets site (GGAA) containing GGAA in the core of the binding site.

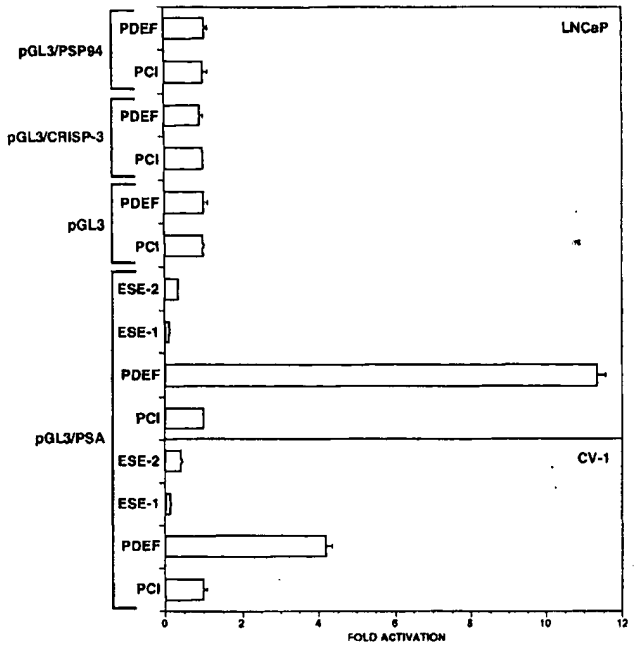


FIG. 5. **Transcriptional activation of the PSA promoter by PDEF.** LNCaP and CV-1 cells were co-transfected with the indicated PDEF, ESE-1, and ESE-2 pCI expression vector constructs or the parental pCI expression vector and luciferase constructs containing the 7-kb PSA promoter, the PSP94 promoter, the CRISP-3 promoter, or the empty pGL3 luciferase vector. Luciferase activity in the lysates was determined 16 h later as described. Data shown are means of triplicate measurements from one representative transfection.

to PDEF ESE-1, ESE-2, ESE-3, NERF-2, ELF-1, and PU.1 bound with high affinity exclusively to the GGAA-containing oligonucleotide, whereas Tel was unable to interact with either of the oligonucleotides (Fig. 4). These results provide strong evidence that the DNA binding specificity of PDEF is distinct from other members of the Ets family and uniquely shifted toward a GGAT recognition sequence.

PDEF Enhances Transcription of the PSA Promoter—To assess whether the PSA gene is indeed a target for PDEF and to determine whether PDEF acts as a repressor or enhancer of transcription, full-length PDEF inserted into the eukaryotic expression vector pCI was co-transfected into PDEF- and AR-negative CV-1 cells together with a reporter gene construct containing the luciferase gene under the control of the 7-kb PSA promoter. Co-transfection with pCI/PDEF resulted in a ~4-fold transcriptional stimulation of the PSA promoter construct compared with the parental pCI vector (Fig. 5). The specificity of this effect was evaluated in co-transfections with other epithelial-restricted members of the Ets family that are expressed in the prostate. Neither ESE-1 nor ESE-2 enhanced PSA promoter activity. On the contrary, ESE-1, and to a lesser extent ESE-2, repressed basal PSA promoter activity (Fig. 5).

To confirm that the PSA promoter is a target for PDEF in prostate epithelium, the same transfection experiment was repeated in LNCaP cells, which are widely used as a model for prostate cancer, since these cells are androgen-sensitive, secrete PSA, and form androgen-sensitive tumors in nude mice (36–41). PDEF enhanced PSA promoter activity 11-fold in LNCaP cells grown with charcoal-stripped serum in the absence of androgen, whereas ESE-1 and ESE-2 repressed the PSA promoter (Fig. 5). The specificity of this transactivation was confirmed in co-transfection experiments with either the empty pGL3 luciferase vector or with the luciferase gene under the control of other promoters (CRISP-3, PSP94) expressed in the prostate. In contrast to the PSA promoter, PDEF was unable to transactivate the CRISP-3 or PSP94 promoter in LNCaP cells (Fig. 5). These results demonstrate that PDEF is a positive regulator of transcription and that the PSA gene is a prostate-specific target for PDEF. Interestingly, PDEF was able to enhance PSA promoter activity independent from androgen, suggesting that PDEF might play a role in androgen-independent PSA gene expression in advanced prostate cancer.

PDEF Interacts with the AR DNA Binding Domain and Enhances Androgen-mediated Transactivation of the PSA Promoter—The AR binding sites in the PSA gene have been shown to be critical for androgen inducibility, and both PDEF and AR are co-expressed in the same cells. A characteristic feature of all Ets factors is their ability to interact with transcription factors of other gene families, suggesting that interaction of PDEF with factors binding to other regulatory elements within the PSA promoter may be an important mechanism of transcriptional control.

To evaluate whether PDEF cooperates with AR, PDEF and AR expression vectors were co-transfected together with pGL3/PSA into AR-negative CV-1 cells and incubated either in the absence or presence of androgen DHT. PDEF in the absence of androgen enhanced PSA promoter activity ~4-fold. Androgen in the absence of PDEF enhanced PSA promoter transcription ~27-fold (Fig. 6). In the presence of both PDEF and androgen, a significant synergistic effect of PSA promoter transactivation was observed leading to a ~57-fold activation, indicating that PDEF and AR cooperate in the regulation of the PSA promoter and that PDEF-mediated PSA promoter transcription contains both an androgen-independent and androgen-dependent component.

To evaluate whether functional interaction between PDEF and AR correlates with physical interaction, GST pull-down experiments were performed with GST fusion proteins containing different domains of AR and *in vitro* translated [^{35}S]methionine-labeled full-length PDEF. Specific retention of PDEF was only observed, when GST/AR fusion proteins containing the DNA binding domain (amino acids 559–624) were used, but not with GST alone or with other domains of AR (Fig. 7, *A* and *B*). These results strongly suggest that PDEF interacts with the DNA binding domain of AR and that cooperativity between PDEF and AR may be due to direct interaction between these two factors.

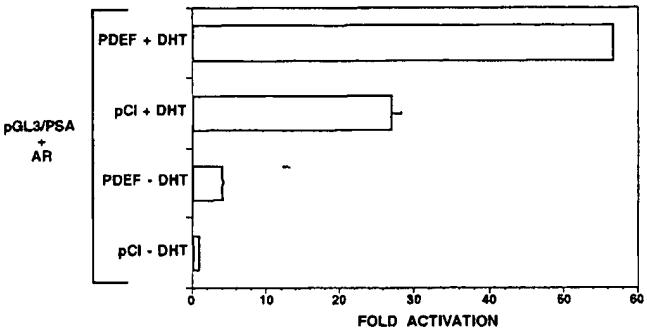


FIG. 6. **PDEF cooperates with AR in transactivation of the PSA promoter.** Figure shows transcriptional activation of the PSA promoter by PDEF in the absence or presence of androgen. CV-1 cells were co-transfected with the indicated PDEF pCI expression vector construct or the parental pCI expression vector together with an AR expression vector and luciferase constructs containing the PSA promoter. Cells were incubated in the absence or presence of DHT.

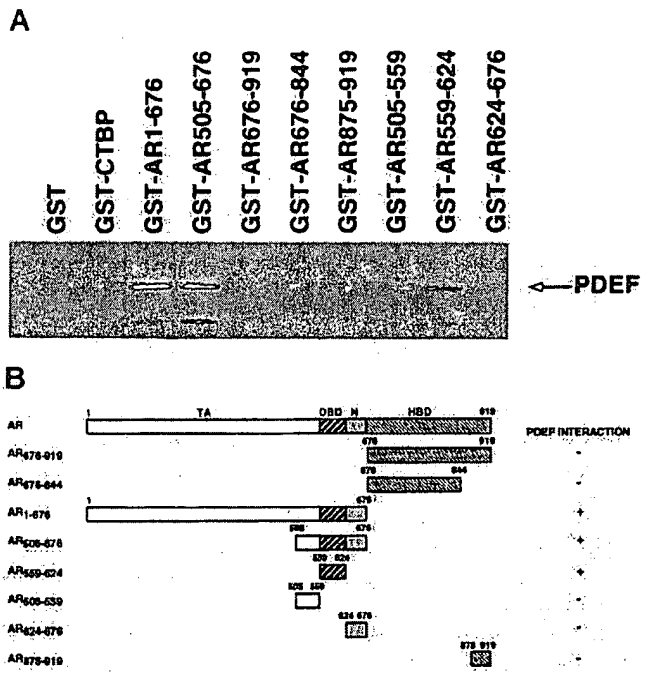


FIG. 7. **PDEF interacts with AR.** *A*, PDEF binds to the AR DNA binding domain in GST pull-down assays. GST fusion proteins with different domains of AR are indicated above Fig. 4*A* and schematically in Fig. 4*B*. *B*, schematic representation of the AR fragments used as GST fusion proteins and of the GST pull-down results with PDEF. Functional domains of the AR are denoted *above* the diagram of full-length AR.

DISCUSSION

Normal function of the prostate gland requires proper differentiation of epithelial cells along a tightly controlled developmental pathway. This process is regulated by distinct sets of transcription factors, leading to a strategically timed switch on or off of specific sets of genes. Deregulation of this pathway leads to aberrant differentiation, immortalization, uncontrolled proliferation, and eventually prostate cancer formation. While some aspects of prostate epithelium-specific gene expression have been elucidated, only a limited number of prostate epithelial cell-restricted transcriptional regulators have been characterized. The Ets transcription factor family plays a particularly important role in epithelial cells, as demonstrated by the isolation of three highly related epithelial-specific Ets factors, ESE-1, ESE-2, and ESE-3 (13–21). Each of these Ets factors, nevertheless, expresses unique expression patterns as well as distinct functions. PDEF is the factor with the most restricted expression pattern, and, whereas ESE-1, ESE-2, and ESE-3 are all members of the same subclass of Ets factors, PDEF is the sole mammalian member of its subclass. The sequence of the PDEF Ets DNA binding domain is significantly different from all other members of the Ets family, which is also reflected in its distinct DNA binding specificity. EMSA analysis of potential Ets binding sites within the PSA promoter demonstrated a significant preference for binding to sites containing a GGAT core rather than the preferred GGAA core typical for other members of the Ets family (12). Direct comparison of the relative DNA binding affinity of PDEF to other members of the Ets family using two oligonucleotides differing only in a single nucleotide change (GGAT *versus* GGAA) confirmed that PDEF strongly prefers interaction with GGAT containing binding sites, whereas other Ets factors such as ESE-1, ESE-2, ESE-3, NERF-2, ELF-1, and PU.1 have a striking preference for GGAA. This DNA binding specificity of PDEF is unique among the Ets family, although some Ets factors have been shown to interact with lower affinity with GGAT binding sites (19, 42). The distinct DNA binding specificity of PDEF also predicts that PDEF recognizes regulatory elements that are distinct from binding sites for other Ets factors, which is in contrast to the majority of Ets factors which all recognize very similar DNA sequences. The distinct DNA binding specificity of PDEF might be due to amino acid differences within helix 3 of the Ets DNA binding domain (Fig. 1*B*) (12). Helix 3 in PU.1, SAP-1, and Ets-1 has previously been shown to directly interact with DNA, and PDEF contains several unique amino acid substitutions that may influence DNA binding specificity (12). Further support for the difference between PDEF activity and other Ets factors is provided by the transactivation experiments demonstrating that PDEF, but not other Ets factors can transactivate the PSA promoter. ESE-1 and to lesser extent ESE-2 actually repress the PSA promoter. Since both PDEF and ESE-1 are expressed in prostate epithelium, PDEF and ESE-1 may play opposing roles in the regulation of prostate epithelium-specific gene expression. Distinct protein-protein interactions, subcellular location, and phosphorylation may modulate the activity of PDEF and ESE-1 and ultimately determine whether a gene is positively or negatively regulated by PDEF and ESE-1. Indeed, both PDEF and ESE-1 contain putative MAP kinase phosphorylation sites as well as sites for other types of kinases. In PDEF, the majority of MAP kinase phosphorylation sites are clustered at the amino terminus. Similar clusters of MAP kinase phosphorylation sites have been observed in the transactivation domains of other Ets factors such as ELK-1 and SAP-1, where they play a critical role in the regulation of transactivation capacity of ELK-1 and SAP-1 (10).

Northern blot analysis indicates the presence of high levels of PDEF transcripts almost exclusively in the prostate and significantly lower amounts in other hormone regulated tissues such as mammary gland, salivary gland, and ovary. Analysis of the public Est data base revealed in addition to prostate multiple entries for PDEF in libraries derived from breast, ovarian, lung, colon, and uterine cancer as well as oligodendroglioma, indicating that PDEF might be overexpressed in several types of cancer including from tissues whose normal counterparts do not express PDEF.

Our *in situ* hybridization and PCR experiments demonstrate that PDEF expression is restricted to the luminal epithelial cells within the prostate as well as LNCaP prostate cancer cells, the same cells that express PSA and AR. This unique expression pattern of PDEF indicates a function for PDEF in prostate gland development and possibly prostate cancer, and,

furthermore, suggests a connection between PDEF and androgen. The level of PSA expression is being widely used as a marker for prostate cancer, and its expression is restricted to epithelial cells of the prostate. The PSA gene is activated in the presence of androgen in normal prostate epithelium as well as in early stages of prostate cancer, and androgen ablation therapy is used at these stages to control prostate cancer growth. Therapeutic efficacy of androgen ablation-mediated prostate cancer regression is being monitored by the decrease in PSA expression, and it is the eventual rise of PSA gene expression that indicates the conversion of a previously androgen-sensitive prostate cancer to an androgen-resistant prostate cancer. Regulation of PSA gene expression in the androgen-sensitive phase of prostate cancer is at least partially regulated by androgen via several AR binding sites and androgen-responsive regions in both the proximal promoter and an upstream enhancer between −3738 and −5824 (7–9, 43). Several putative Ets sites are located in proximity to AR binding sites in both the promoter and enhancer region. At least two of the PDEF binding sites (sites C and E) are within the enhancer region, whereas two other PDEF binding sites (sites G and H) are between the promoter and enhancer region within a poorly defined region of the PSA gene.

Activation of PSA gene expression in advanced prostate cancer upon conversion to androgen resistance suggests an androgen-independent regulatory mechanism. The mechanisms and the transcription factors involved in androgen-independent PSA gene expression have not been characterized in detail, although several possible mechanisms such as AR amplifications, AR mutations, androgen-independent AR activation due to interaction with co-activators or phosphorylation have been proposed (44–46). The location of PDEF specifically in prostate epithelium and the ability of PDEF to bind to and transactivate the PSA promoter in the absence of androgen as well as to enhance androgen mediated PSA gene expression suggest that deregulation of PDEF activity may be one mechanism in prostate cancer to induce androgen-independent PSA gene expression.

The interaction of AR with various co-activators has been demonstrated to be a critical step in its transcriptional activity and many times involves the interaction of the sequence LXXLL with the carboxyl-terminal AF-2 domain of AR (47). PDEF indeed contains this sequence within the Ets DNA binding domain. Nevertheless, we do not detect any direct interaction between PDEF and the AF-2 domain, but instead with the DNA binding domain of the AR. Although most co-regulators interact with either the amino-terminal transactivation domain or the carboxyl-terminal ligand binding domain and AF-2 domain, a few factors such as SNURF and c-Jun have been shown to interact with the DNA binding domain of AR (48–50). Nevertheless, whereas the AR DNA binding domain is necessary but not sufficient for interaction with SNURF, PDEF interacts with the DNA binding domain alone. Like PDEF, SNURF or c-Jun enhance androgen-induced transactivation, but also act as AR-independent transactivators. However, in contrast to most other co-regulators, which are ubiquitously expressed and, thus, presumably act as co-regulators for a number of other transcription factors as well, PDEF is in a unique position being primarily expressed in the same cells that express the AR.

Although some of the high affinity binding sites for PDEF are in proximity to AR binding sites, the distances between these sites except for sites I and J appear to be too big to allow protein-protein interactions without some form of looping. However, as has been shown for a variety of genes, transcriptional activation of genes requires the assembly of an enhanceosome, consisting of various transcriptional activators and repressors as well as architectural proteins that form a multiprotein complex via long distances (51). Enhanceosome assembly involves both allosteric changes of DNA leading to looping as well as protein-protein interactions (51). In addition, some transcription factors can form protein-protein interactions with other transcription factors in the absence of DNA binding. It is thus conceivable that PDEF interacts with AR without binding to DNA itself. Ultimate transcriptional activation apparently involves the interaction of various co-activators and/or co-repressors with a whole set of transcription factors on one hand and with the basal transcription complex on the other hand. These co-activators act like a scaffold protein with a whole set of different surfaces interacting at different sites with a large number of transcription factors. PDEF might be one additional transcription factor that recruits co-activators and by interaction with AR in a combinatorial manner enhances the activity of the co-activator, leading to enhanced transcription.

In conclusion, our results provide evidence that PDEF is a novel prostate epithelium-specific transcription factor and co-regulator of the AR that can act both synergistically with the AR as well as androgen-independent in enhancing PSA gene expression. Our data support the notion that PDEF function may exemplify one mechanism of androgen-independence upon prostate cancer progression and the concomitant reactivation of PSA gene expression.

Acknowledgments—We acknowledge fruitful discussions with Dr. Ellen Gravalese, Dr. Glen Bubley, and Dr. Steve Balk.

REFERENCES

1. Small, E. J. (1999) *Curr. Opin. Oncol.* **11**, 226–235
2. Nupponen, N. & Visakorpi, T. (1999) *Eur. Urol.* **35**, 351–354
3. Rhim, J. S., Webber, M. M., Bello, D., Lee, M. S., Arnstein, P., Chen, L. S. & Jay, G. (1994) *Proc. Natl. Acad. Sci. U. S. A.* **91**, 11874–11878
4. Robinson, E. J., Neal, D. E. & Collins, A. T. (1998) *Prostate* **37**, 149–160
5. Ablin, R. J. (1997) *J. Cancer Res. Clin. Oncol.* **123**, 583–594
6. Sadar, M. D. (1999) *J. Biol. Chem.* **274**, 7777–7783
7. Cleutjens, K. B., van der Korput, H. A., Ehren-van Eekelen, C. C., Sikes, R. A., Fasciana, C., Chung, L. W. & Trapman, J. (1997) *Mol. Endocrinol.* **11**, 1256–1265
8. Cleutjens, K. B., van der Korput, H. A., van Eekelen, C. C., van Rooij, H. C., Faber, P. W. & Trapman, J. (1997) *Mol. Endocrinol.* **11**, 148–161
9. Zhang, S., Murtha, P. E. & Young, C. Y. (1997) *Biochem. Biophys. Res. Commun.* **231**, 784–788
10. Wasylyk, B., Hagman, J. & Gutierrez-Hartmann, A. (1998) *Trends Biochem. Sci.* **23**, 213–216
11. Dittmer, J. & Nordheim, A. (1998) *Biochim. Biophys. Acta* **1377**, F1–F11
12. Graves, B. J. & Petersen, J. M. (1998) *Adv. Cancer Res.* **75**, 1–55
13. Oettgen, P., Barcinski, M., Boltax, J., Stolt, P., Akbarali, Y. & Libermann, T. A. (1999) *Genomics* **55**, 358–362
14. Oettgen, P., Carter, K. C., Augustus, M., Barcinski, M., Boltax, J., Kunsch, C. & Libermann, T. A. (1997) *Genomics* **445**, 456–457
15. Oettgen, P., Alani, R. M., Barcinski, M. A., Brown, L., Akbarali, Y., Boltax, J., Kunsch, C., Munger, K. & Libermann, T. A. (1997) *Mol. Cell. Biol.* **17**, 4419–4433
16. Andreoli, J. M., Jang, S. I., Chung, E., Coticchia, C. M., Steinert, P. M. & Markova, N. G. (1997) *Nucleic Acids Res.* **25**, 4287–4295
17. Zhou, J., Ng, A. Y., Tymms, M. J., Jermiin, L. S., Seth, A. K., Thomas, R. S. & Kola, I. (1998) *Oncogene* **17**, 2719–2732
18. Neve, R., Chang, C. H., Scott, G. K., Wong, A., Friis, R. R., Hynes, N. E. & Benz, C. C. (1998) *FASEB J.* **12**, 1541–1550
19. Tymms, M. J., Ng, A. Y., Thomas, R. S., Schutte, B. C., Zhou, J., Eyre, H. J., Sutherland, G. R., Seth, A., Rosenberg, M., Papas, T., Debouck, C. & Kola, I. (1997) *Oncogene* **15**, 2449–2462
20. Chang, C. H., Scott, G. K., Kuo, W. L., Xiong, X., Suzdaltseva, Y., Park, J. W., Sayre, P., Erny, K., Collins, C., Gray, J. W. & Benz, C. C. (1997) *Oncogene* **14**, 1617–1622
21. Oettgen, P., Kas, K., Dube, A., Gu, X., Grall, F., Thamrongsak, U., Akbarali, Y., Finger, E., Boltax, J., Endress, G., Munger, K., Kunsch, C. & Libermann, T. A. (1999) *J. Biol. Chem.* **274**, 29439–29452
22. Libermann, T. A. & Baltimore, D. (1993) *Mol. Cell. Biol.* **13**, 5957–5969
23. Oettgen, P., Akbarali, Y., Boltax, J., Best, J., Kunsch, C. & Libermann, T. A. (1996) *Mol. Cell. Biol.* **16**, 5091–5106
24. Libermann, T. A., Friesel, R., Jaye, M., Lyall, R. M., Westermark, B., Drohan, W., Schmidt, A., Maciag, T., Schlessinger, J. (1987) *EMBO J.* **6**, 1627–1632
25. Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. (1979) *Biochemistry* **18**, 5294–5299
26. Lopez, M., Oettgen, P., Akbarali, Y., Dendorfer, U. & Libermann, T. A. (1994) *Mol. Cell. Biol.* **14**, 3292–3309
27. Brinkmann, A. O., Faber, P. W., van Rooij, H. C., Kuiper, G. G., Ris, C., Klaassen, P., van der Korput, J. A., Voorhorst, M. M., van Laar, J. H.,

Mulder, E., and Trapman, J. (1989) *J. Steroid Biochem.* **34,** 307–310
28. Pahl, H. L., Scheibe, R. J., Zhang, D. E., Chen, H. M., Galson, D. L., Maki, R. A. & Tenen, D. G. (1993) *J. Biol. Chem.* **268,** 5014–5020
29. Farr, A. & Roman, A. (1992) *Nucleic Acids Res.* **20,** 920
30. Sun, Z., Pan, J. & Balk, S. P. (1997) *Nucleic Acids Res.* **25,** 3318–3325
31. Smith, D. B. & Johnson, K. S. (1988) *Gene (Amst.)* **67,** 31–40
32. Songyang, Z., Blechner, S., Hoagland, N., Hoekstra, M. F., Piwnica, W. H. & Cantley, L. C. (1994) *Curr. Biol.* **4,** 973–982
33. Wei, Z., Angerer, R. C. & Angerer, L. M. (1999) *Mol. Cell. Biol.* **19,** 1271–1278
34. Chen, T., Bunting, M., Karim, F. D. & Thummel, C. S. (1992) *Dev. Biol.* **151,** 176–191
35. Stapleton, D., Balan, I., Pawson, T. & Sicheri, F. (1999) *Nat. Struct. Biol.* **6,** 44–49
36. Chung, L. W., Zhau, H. E. & Wu, T. T. (1997) *J. Cell. Biochem. Suppl.* **28–29,** 174–181
37. Chung, L. W., Li, W., Gleave, M. E., Hsieh, J. T., Wu, H. C., Sikes, R. A., Zhau, H. E., Bandyk, M. G., Logothetis, C. J., Rubin, J. S., and Voneschenbach, A. C. (1992) *J. Cell. Biochem. Suppl.* **16H,** 99–105
38. Wang, X., An, Z., Geller, J. & Hoffman, R. M. (1999) *Prostate* **39,** 182–186
39. Mehta, P. P., Perez-Stable, C., Nadji, M., Mian, M., Asotra, K. & Roos, B. A. (1999) *Dev. Genet.* **24,** 91–110
40. Chen, C. T., Gan, Y., Au, J. L. & Wientjes, M. G. (1998) *Cancer Res.* **58,** 2777–2783
41. Chen, Y., Martinez, L. A., LaCava, M., Coghlan, L. & Conti, C. J. (1998) *Oncogene* **16,** 1913–1920
42. Bosselut, R., Levin, J., Adjadj, E. & Ghysdael, J. (1993) *Nucleic Acids Res.* **21,** 5184–5191
43. Huang, W., Shostak, Y., Tarr, P., Sawyers, C. & Carey, M. (1999) *J. Biol. Chem.* **274,** 25756–25768
44. Yeh, S., Lin, H. K., Kang, H. Y., Thin, T. H., Lin, M. F. & Chang, C. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 5458–5463
45. Culig, Z., Hobisch, A., Hittmair, A., Cronauer, M. V., Radmayr, C., Bartsch, G. & Klocker, H. (1997) *Drugs Aging* **10,** 50–58
46. Koivisto, P., Kononen, J., Palmberg, C., Tammela, T., Hyytinen, E., Isola, J., Trapman, J., Cleutjens, K., Noordzij, A., Visakorpi, T. & Kallioniemi, O. P. (1997) *Cancer Res.* **57,** 314–319
47. McInerney, E. M., Rose, D. W., Flynn, S. E., Westin, S., Mullen, T. M., Krones, A., Inostroza, J., Torchia, J., Nolte, R. T., Assa-Munt, N., Milburn, M. V., Glass, C. K. & Rosenfeld, M. G. (1998) *Genes Dev.* **12,** 3357–3368
48. Bubulya, A., Wise, S. C., Shen, X. Q., Burmeister, L. A. & Shemshedini, L. (1996) *J. Biol. Chem.* **271,** 24583–24589
49. Moilanen, A. M., Poukka, H., Karvonen, U., Hakli, M., Janne, O. A. & Palvimo, J. J. (1998) *Mol. Cell. Biol.* **18,** 5128–5139
50. Aarnisalo, P., Palvimo, J. J. & Janne, O. A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 2122–2127
51. Yie, J., Merika, M., Munshi, N., Chen, G. & Thanos, D. (1999) *EMBO J.* **18,** 3074–3089

SeqServer

BLAST2 Search Results

Sequences | Help

Retrieval | BLAST2 | MSA | Clustal W | CAGAssembly | Fmap | Translation | BLAST2 Manual

Program: blastp
Sequence ID(s):
1813005CD1 vs. genpept137

NCBI-BLASTP 2.0.10 [Aug-26-1999]

Reference: Altschul, Stephen F., Thomas L. Madden, Alejandro A. Schaffer, Jinghui Zhang, Zheng Zhang, Webb Miller, and David J. Lipman (1997), "Gapped BLAST and PSI-BLAST: a new generation of protein database search programs", Nucleic Acids Res. 25:3389-3402.

Query= 1813005CD1
(335 letters)

Database: genpept137
1,534,369 sequences; 474,463,515 total letters

Searching.....done

Sequences producing significant alignments:	Score (bits)	E Value
g6721498 prostate ets [Homo sapiens]	699	0.0
g1007418 Ets transcription factor PDEF [Homo sapiens]	699	0.0
g18204266 prostate epithelium-specific Ets transcription fact	699	0.0
g14598856 unnamed protein product [Homo sapiens]	699	0.0
g11876976 unnamed protein product [Homo sapiens]	699	0.0
g11876968 unnamed protein product [Homo sapiens]	699	0.0
g4760537 Prostate specific transcription factor ets [Mus musc	598	e-169
g26347555 unnamed protein product [Mus musculus]	598	e-169
g15215055 Pse-pending protein [Mus musculus]	598	e-169
g4406284 Ets4 transcription factor [Strongylocentrotus purpur	175	2e-42

>g6721498 prostate ets [Homo sapiens]
Length = 335

Score = 699 bits (1785), Expect = 0.0
Identities = 335/335 (100%), Positives = 335/335 (100%)

```
Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60
           MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL
Sbjct: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE
Sbjct: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG
Sbjct: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV
Sbjct: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM
Sbjct: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI
Sbjct: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
```

>g1007418 Ets transcription factor PDEF [Homo sapiens]
Length = 335

Score = 699 bits (1785), Expect = 0.0
Identities = 335/335 (100%), Positives = 335/335 (100%)

```
Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60
           MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL
Sbjct: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE
Sbjct: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG
Sbjct: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV
Sbjct: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM
Sbjct: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI
Sbjct: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
```


>g18204266 prostate epithelium-specific Ets transcription factor [Homo sapiens]
Length = 335

Score = 699 bits (1785), Expect = 0.0
Identities = 335/335 (100%), Positives = 335/335 (100%)

```
Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60
           MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL
Sbjct: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE
Sbjct: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG
Sbjct: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV
Sbjct: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM
Sbjct: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI
Sbjct: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
```

>g14598856 unnamed protein product [Homo sapiens]
Length = 335

Score = 699 bits (1785), Expect = 0.0
Identities = 335/335 (100%), Positives = 335/335 (100%)

```
Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60
           MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL
Sbjct: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE
Sbjct: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG
Sbjct: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV
Sbjct: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM
Sbjct: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI
Sbjct: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
```


>g11876976 unnamed protein product [Homo sapiens]
Length = 335

Score = 699 bits (1785), Expect = 0.0
Identities = 335/335 (100%), Positives = 335/335 (100%)

```
Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60
           MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL
Sbjct: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE
Sbjct: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG
Sbjct: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV
Sbjct: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM
Sbjct: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI
Sbjct: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
```

>g11876968 unnamed protein product [Homo sapiens]
Length = 335

Score = 699 bits (1785), Expect = 0.0
Identities = 335/335 (100%), Positives = 335/335 (100%)

```
Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60
           MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL
Sbjct: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQGLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE
Sbjct: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG
Sbjct: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV
Sbjct: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
```

```
          DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM
Sbjct: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI
Sbjct: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
```

>g4760537 Prostate specific transcription factor ets [Mus musculus]
Length = 325

```
 Score =  598 bits (1524), Expect = e-169
 Identities = 287/335 (85%), Positives = 302/335 (89%), Gaps = 10/335 (2%)

Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQQLSAFYL 60
           MGSASPGLS+VSP  LLL PD   RTG EKAA+GA+G E+++WSPSPPATPEQQLSAFYL
Sbjct: 1   MGSASPGLSNVSPGCLLLFPDVAPRTGTEKAASGAMGPEKQEWSPSPPATPEQQLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYF+M YP+DSSW AK P A + E+ PEEPEQCPVIDSQA          G TL+EHSLE
Sbjct: 61  SYFNM-YPDDSSWVAKVPEARAGEDHPEEPEQCPVIDSQA---------SGSTLDEHSLE 110

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADP DWSP NVQKWLLWTEHQYRLPP GKAFQEL G
Sbjct: 111 QVQSMVVGEVLKDIETACKLLNITADPGDWSPGNVQKWLLWTEHQYRLPPAGKAFQELGG 170

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQR+PLGGDVLHAHLDIWKSAAWMKERTSPG +HYCASTSEE WTD EV
Sbjct: 171 KELCAMSEEQFRQRAPLGGDVLHAHLDIWKSAAWMKERTSPGTLHYCASTSEEGWTDGEV 230

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWG+RKNRPAM
Sbjct: 231 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGVRKNRPAM 290

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP+
Sbjct: 291 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPV 325
```

>g26347555 unnamed protein product [Mus musculus]
Length = 325

```
 Score =  598 bits (1524), Expect = e-169
 Identities = 287/335 (85%), Positives = 302/335 (89%), Gaps = 10/335 (2%)

Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQQLSAFYL 60
           MGSASPGLS+VSP  LLL PD   RTG EKAA+GA+G E+++WSPSPPATPEQQLSAFYL
Sbjct: 1   MGSASPGLSNVSPGCLLLFPDVAPRTGTEKAASGAMGPEKQEWSPSPPATPEQQLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYF+M YP+DSSW AK P A + E+ PEEPEQCPVIDSQA          G TL+EHSLE
Sbjct: 61  SYFNM-YPDDSSWVAKVPEARAGEDHPEEPEQCPVIDSQA---------SGSTLDEHSLE 110

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADP DWSP NVQKWLLWTEHQYRLPP GKAFQEL G
Sbjct: 111 QVQSMVVGEVLKDIETACKLLNITADPGDWSPGNVQKWLLWTEHQYRLPPAGKAFQELGG 170

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQR+PLGGDVLHAHLDIWKSAAWMKERTSPG +HYCASTSEE WTD EV
Sbjct: 171 KELCAMSEEQFRQRAPLGGDVLHAHLDIWKSAAWMKERTSPGTLHYCASTSEEGWTDGEV 230

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWG+RKNRPAM
Sbjct: 231 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGVRKNRPAM 290

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP+
Sbjct: 291 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPV 325
```



```
Sbjct: 171 KELCAMSEEQFRQRAPLGGDVLHAHLDIWKSAAWMKERTSPGTLHYCASTSEEGWTDGEV 230

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWG+RKNRPAM
Sbjct: 231 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGVRKNRPAM 290

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP+
Sbjct: 291 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPV 325
```

>g15215055 Pse-pending protein [Mus musculus]
Length = 325

```
 Score =  598 bits (1524), Expect = e-169
 Identities = 287/335 (85%), Positives = 302/335 (89%), Gaps = 10/335 (2%)

Query: 1   MGSASPGLSSVSPSHLLLPPDTVSRTGLEKAAAGAVGLERRDWSPSPPATPEQQLSAFYL 60
           MGSASPGLS+VSP  LLL PD   RTG EKAA+GA+G E+++WSPSPPATPEQQLSAFYL
Sbjct: 1   MGSASPGLSNVSPGCLLLFPDVAPRTGTEKAASGAMGPEKQEWSPSPPATPEQQLSAFYL 60

Query: 61  SYFDMLYPEDSSWAAKAPGASSREEPPEEPEQCPVIDSQAPAGSLDLVPGGLTLEEHSLE 120
           SYF+M YP+DSSW AK P A + E+ PEEPEQCPVIDSQA          G TL+EHSLE
Sbjct: 61  SYFNM-YPDDSSWVAKVPEARAGEDHPEEPEQCPVIDSQA---------SGSTLDEHSLE 110

Query: 121 QVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWTEHQYRLPPMGKAFQELAG 180
           QVQSMVVGEVLKDIETACKLLNITADP DWSP NVQKWLLWTEHQYRLPP GKAFQEL G
Sbjct: 111 QVQSMVVGEVLKDIETACKLLNITADPGDWSPGNVQKWLLWTEHQYRLPPAGKAFQELGG 170

Query: 181 KELCAMSEEQFRQRSPLGGDVLHAHLDIWKSAAWMKERTSPGAIHYCASTSEESWTDSEV 240
           KELCAMSEEQFRQR+PLGGDVLHAHLDIWKSAAWMKERTSPG +HYCASTSEE WTD EV
Sbjct: 171 KELCAMSEEQFRQRAPLGGDVLHAHLDIWKSAAWMKERTSPGTLHYCASTSEEGWTDGEV 230

Query: 241 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGIRKNRPAM 300
           DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWG+RKNRPAM
Sbjct: 231 DSSCSGQPIHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLWGVRKNRPAM 290

Query: 301 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPI 335
           NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP+
Sbjct: 291 NYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHPV 325
```

>g4406284 Ets4 transcription factor [Strongylocentrotus purpuratus]
Length = 363

```
 Score =  175 bits (439), Expect = 2e-42
 Identities = 104/283 (36%), Positives = 146/283 (50%), Gaps = 51/283 (18%)

Query: 103 GSLDLVPGGLTLEEHSLEQVQSMVVGEVLKDIETACKLLNITADPMDWSPSNVQKWLLWT 162
           GS     GG  + E    ++  +++ E++ DI   C  LN+  +   W+P +VQKWLL
Sbjct: 80  GSTIQQQQGNQISEIEENEMMRLLMREMIPDILEDCHKLNLIPNVYCWTPEDVQKWLLCV 139

Query: 163 EHQYRLPPMGKAFQELAGKELCAMSEEQFRQRSPLGGDVLHAHLDIWKSA---------- 212
            +++ L  +      + G  L  + +  FR R+P  GD+L++ + + + KS+
Sbjct: 140 ANRFELGBLEMGHFYINGPTLATLQDVDFRHRAPKCGDILYSVVCLLKSSIHFDVAPQVV 199

Query: 213 --AWMKERTSPGAIHYCA----------------------------STSEESWTDSEVDS 242
             A     P  H+                               ST  S  +  + S
Sbjct: 200 PQAQGMMPNIPQPSHFITGAPTTTMMPPKYTYTPVSPVATPGDESGSTGSPSDDEISIPS 259
```



```
Query: 243 SCSGQP----------IHLWQFLKELLLKPHSYGRFIRWLNKEKGIFKIEDSAQVARLW 291
                P           I LWQFLKELLL+P+SY   IRW+++ +GIFKIEDS +VARLW
Sbjct: 260 PAPSGPNTSMTPNHTGGIQLWQFLKELLLQPNSYSYCIRWIDRTQGIFKIEDSVEVARLW 319

Query: 292 GIRKNRPAMNYDKLSRSIRQYYKKGIIRKPDISQRLVYQFVHP 334
           G+RKNRPAMNYDKLSRSIRQYY+KGI++K ++SQRLVYQFV P
Sbjct: 320 GLRKNRPAMNYDKLSRSIRQYYRKGIMKKTEVSQRLVYQFVTP 362
```

```
  Database: genpept137
    Posted date:  Sep 11, 2003 11:22 AM
  Number of letters in database: 474,463,515
  Number of sequences in database:  1,534,369

Lambda     K      H
   0.315   0.132    0.410

Gapped
Lambda     K      H
   0.270   0.0470    0.230

Matrix: BLOSUM62
Gap Penalties: Existence: 11, Extension: 1
Number of Hits to DB: 357701212
Number of Sequences: 1534369
Number of extensions: 15704775
Number of successful extensions: 40279
Number of sequences better than 10.0: 412
Number of HSP's better than 10.0 without gapping: 362
Number of HSP's successfully gapped in prelim test: 50
Number of HSP's that attempted gapping in prelim test: 39494
Number of HSP's gapped (non-prelim): 556
length of query: 335
length of database: 474,463,515
effective HSP length: 56
effective length of query: 279
effective length of database: 388,538,851
effective search space: 108402339429
effective search space used: 108402339429
T: 11
A: 40
X1: 16 ( 7.3 bits)
X2: 38 (14.8 bits)
X3: 64 (24.9 bits)
S1: 41 (21.6 bits)
```

Submit sequences to: BLAST2 Submit Reset